



**FACULTAD DE CIENCIAS (Departamento de Biología Molecular)**

**TESIS DOCTORAL:**

**“Análisis estructural y reconocimiento de ligandos (no integrinas) en la familia  
ICAM de receptores celulares”**

**Autor:**

**David Jiménez Gil**

**Director:**

**Dr. José M. Casasnovas Suelves  
Departamento de Estructura de Macromoléculas  
CNB/CSIC**

**Madrid, Mayo de 2007**

*A mis padres*

## **AGRADECIMIENTOS**

Me gustaría darle las gracias en primer lugar a mi director de tesis el Dr. José María Casanovas por haberme dado la oportunidad de llevar a cabo este proyecto, por sus consejos y su inestimable ayuda.

Una de las mejores cosas que tiene este trabajo es la cantidad y calidad de la gente que se conoce, por eso quiero daros las gracias a todos los que habéis compartido las largas horas de trabajo y también de diversión conmigo a lo largo de tantos años, empezando por todos los amigos de Salamanca, especialmente Pedro y Raúl, que estaba convencido de que acabaría haciendo la tesis. Tampoco me quiero olvidar de los chicos de Suecia, Alberto, Yolanda, Ángel G, Ángel C, Nilla, Antonio, Cristina, Angelita, Jose, Amilcar, Rafa, Birgitta, Ghasem, Sam, Marcos y muy especialmente a mi primo Luís Miguel por hacerme sentir en casa estando a miles de kilómetros.

Gracias a todos mis compañeros del CNB, empezando por Desi, que fue el primero a bordo y con el que pasé más horas durante el primer año en Madrid, y a todos los que habéis pasado por el laboratorio, Beatriz, Marta, Olatz, Ángela y Román.

Quiero darle las gracias a Pedro Roda por su ayuda durante su estancia con nosotros en el Karolinska, pero sobre todo por ser un buen amigo y por los buenos ratos que hemos pasado junto con su inseparable Majonie y el resto de tropa de La Princesa. Gracias también a los chicos del 317, Javier, Julián, los “Luises”, Manolo e Irene, por esas cañitas que espero podamos repetir cuando por fin nos toque el Euromillón.

Gracias a todos los amigos que me acogieron cuando volví de Suecia, Sonia, Juan Luis, Vero, Héctor, el Pichón, y todo el resto de chavales.

Finalmente quiero darle las gracias muy especialmente a César, por ser una de las mejores personas que conozco, y el mejor compañero de trabajo que se puede tener, siempre dispuesto a dejar lo que esté haciendo para ayudar a los demás. Por todos los ratos buenos que hemos pasado juntos, porque ratos malos no recuerdo ninguno.

También quiero darle las gracias a mi familia por su ayuda, su paciencia y su comprensión durante todo este tiempo. A mis padres por la ilusión que les hace esta tesis. A Merce y Jose por todo el tiempo que me han aguantado estos últimos años. A mis abuelos, y muy especialmente a tío Satur y a tía Mari, por hacerme sentir en casa cuando empecé a vivir en Madrid.

Por encima de todo gracias a Anabel, por ser todo lo que siempre he deseado, por su optimismo, su valentía y por ser capaz de despertarme todos los días con una sonrisa en la boca.



---

**SUMMARY: Structural analysis and non-integrin ligand recognition in the ICAM (intercellular adhesion molecules) receptor family.**

Intercellular adhesion molecules (ICAM) are heavily glycosylated cell surface receptors of the immunoglobulin superfamily (IgSF). They share functional and structural homologies and mediate cell-cell adhesion interactions relevant for the function of the immune system. Adhesive interactions between members of the ICAM subfamily and leukocyte integrins LFA-1 and Mac-1 mediate leukocyte trafficking through inflamed and uninfamed tissues and contribute to antigen-specific T-cell responses. The crystal structures of the glycosylated N-terminal two domains of ICAM-1 and ICAM-2 provided a framework for understanding the role of glycosylation in the structure and function of the ICAMs. Conserved glycans were less flexible in the structures, interacting with protein residues and contributing to receptor folding and expression. The first *N*-linked glycan in ICAM-2 contacted an exposed tryptophan residue and defined a conserved motif critical for the conformation of the N-terminal, integrin binding domain.

The first domain of ICAM-1 is recognized by the LFA-1 integrin and non-integrin ligands, such as human rhinoviruses and *Plasmodium falciparum* derived proteins. While soluble molecules of ICAM-1 with the two N-terminal IgSF domains are functional in ligand binding, expression of a single N-terminal binding domain has been unsuccessful. We expressed a single N-terminal domain of ICAM-1 by introducing *N*-linked glycosylation in different location of the molecules and observed that domain expression depends of the location of the glycans. Experiments with soluble molecules having the N-terminal two domains of human ICAMs identified glycans of the high mannose type *N*-linked to the second domain of ICAM-2 and ICAM-3, absent in ICAM-1. About 40% of those receptor molecules bear endoglycosidase H sensitive glycans and they were responsible of the receptors DC-SIGN lectin binding activity. High mannose glycans were absent in ICAM-1, which did not bind to the lectin, but they appeared in ICAM-1 mutants with inserted glycosylations.

ICAM-5 also known as telencephalin is expressed in neurons and, differing from the remaining ICAMs, mediates homophilic adhesion interactions through its N-terminal 5 domains. With the purpose of characterizing these interactions we made crystals with soluble ICAM-5 variants and solved three structures of a 4 domain molecule. The structure showed a hockey stick molecular shape and suggested a model for the homophilic interaction, where dimeric ICAM-5 molecules from opposite cells interact in *trans* through complementary charged regions on the N and C-terminal domain modules.



## **CAPITULO 1: INTRODUCCIÓN**



## 1. INTRODUCCIÓN

### 1.1. Adhesión celular

Los procesos de adhesión celular son fundamentales para la organización, comunicación, y motilidad celular en los organismos pluricelulares. Estos procesos están mediados por moléculas de adhesión localizadas en la membrana celular. Las interacciones de moléculas de adhesión con las macromoléculas que constituyen la matriz extracelular constituyen las bases para el mantenimiento de la estructura tisular. Asimismo, las interacciones de la región intracelular de algunos receptores celulares con el citoesqueleto o con proteínas intracelulares de señalización están íntimamente relacionadas con la motilidad celular, así como con procesos de diferenciación y proliferación celular.

Los procesos de adhesión celular podrían dividirse en tres tipos:

1. **Procesos dinámicos:** comprenden adhesiones celulares que participan en migración celular. En ellos las moléculas de adhesión interaccionan entre si durante un corto periodo de tiempo (apartado 1.1.1).
2. **Procesos transitorios:** comprenderían interacciones de adhesión relacionadas con procesos de comunicación celular, que median señales intracelulares implicadas en procesos tan esenciales como respuestas inmunitarias (apartado 1.1.2)
3. **Procesos estáticos:** relacionados con la organización celular en tejidos de organismos pluricelulares (apartado 1.1.3).

#### 1.1.1. Adhesión dinámica, migración celular

Determinados tipos celulares, como los sanguíneos, deben de estar en continuo movimiento ya sea para el transporte de nutrientes o para la detección de cualquier tipo de infección o anomalía, tal y como ocurre con leucocitos o células dendríticas. La migración celular es un proceso complejo que depende del tipo celular que se vea inmerso en el proceso, de su estado de activación y del entorno vascular a lo largo del organismo (Fig.1.1). Los leucocitos migran a través de diversos tejidos siguiendo un

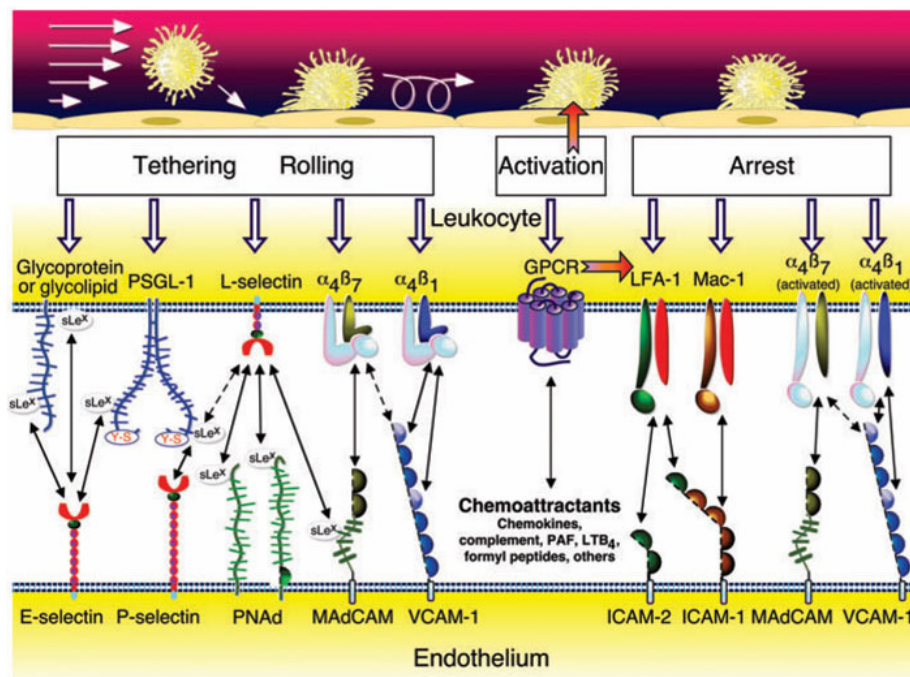
determinado patrón de migración que depende del tipo celular y de su estado de diferenciación y activación. Así, podemos distinguir tres tipos celulares con patrones de migración diferentes (Roitt I 1998):

- **Fagocitos:** incluyendo neutrófilos y monocitos. Dejan la médula espinal y migran hacia los tejidos periféricos donde se encuentra la infección o inflamación. Los neutrófilos hacen un viaje de no retorno, pero los monocitos se diferencian a macrófagos y pueden recircular hacia los tejidos linfáticos secundarios como células presentadoras de antígeno.
- **Linfocitos vírgenes:** migran desde el timo y la médula espinal a los tejidos linfáticos secundarios, principalmente ganglios y bazo. En estos tejidos entran en contacto con antígenos y otros tipos celulares lo que induce su activación y diferenciación. Las células T activadas migran hacia las zonas de inflamación, mientras las células B y las células T de memoria pueblan otros tejidos linfáticos.
- **Células dendríticas:** se originan a partir de células madre de la médula ósea, y colonizan varios órganos. Su papel fundamental es la captura del antígeno. Las células dendríticas migran hacia los ganglios linfáticos locales para presentar el antígeno a las células T CD4<sup>+</sup>.

Las rutas de drenaje linfático y la migración celular aseguran que los linfocitos y las células presentadoras de antígeno converjan en los ganglios linfáticos. La interacción de estos dos tipos celulares conduce a la proliferación de linfocitos específicos en los tejidos linfáticos secundarios antes de ser puestos en circulación. La migración de esas células desde el torrente sanguíneo depende de la expresión de moléculas de adhesión tanto en su superficie como en el endotelio. La expresión de dichas moléculas en tejidos inflamados dirige el tráfico de linfocitos activados y macrófagos hacia esas zonas (Fig.1.1).

Hay dos etapas fundamentales en la migración de los leucocitos a través de los endotelios, proceso complejo en el que participan diferentes moléculas de adhesión celular. La primera etapa del proceso es el anclaje de las células circulantes al endotelio vascular, que está seguida por el movimiento de los leucocitos entre o a través de las células del endotelio (Fig.1.1) (Roitt I 1998). En las etapas iniciales del proceso de adhesión participan las selectinas de leucocitos (selectina-L) y de los endotelios

(selectina –E y –P), que junto con las integrinas  $\beta_4$  estabilizan la unión inicial y reducen la velocidad de los leucocitos, haciéndolos “rodar” por el endotelio (Fig.1.1). La activación de los leucocitos por las quimioquinas genera cambios estructurales en las integrinas leucocitarias, de forma que se incrementa la afinidad por sus ligandos en células endoteliales. Las integrinas más importantes que participan en este proceso son de la subfamilia  $\beta_2$ , especialmente LFA-1 (CD11a-CD18 o  $\alpha_L\beta_2$ ), y  $\beta_4$ ,  $\alpha_4\beta_1$  (VLA) y  $\alpha_4\beta_7$  (Fig.1.1). Estas integrinas se unen a receptores celulares de superficie pertenecientes a la superfamilia de las inmunoglobulinas, como los ICAMs, que se unen a LFA-1, VCAM-1, que es el principal ligando de  $\alpha_4\beta_1$ ; y MAdCAM-1, principal ligando de  $\alpha_4\beta_7$  (Fig.1.1) (Luster AD 2005). Estas interacciones son de mayor afinidad que la de las selectinas y sus ligandos y median la adhesión firme de los leucocitos a las



**Fig.1.1. Migración de leucocitos a través del endotelio.** En la parte superior se indican los pasos de adhesión que han de dar los leucocitos circulantes para atravesar el endotelio. En la parte central se indican las moléculas expresadas por los leucocitos que participan en cada paso y debajo sus respectivos receptores en el endotelio celular. Los leucocitos circulantes en el torrente circulatorio interaccionan con ligandos de las células del endotelio y “ruedan” despacio por la superficie del mismo. Las moléculas de selectina E-, L- y P junto con las integrinas  $\alpha_4$  estabilizan y reducen la velocidad con que ruedan los linfocitos por el endotelio. Cuando una quimioquina específica se une a su receptor sobre una célula “rodante” se genera una señal intracelular de activación para integrinas  $\alpha_2$  y/o  $\alpha_4$ , que cambian de conformación y se unen a los receptores de tipo inmunoglobulina presentes en el endotelio. (Figura adaptada de Luster AD, 2005) (Luster AD 2005).

células endoteliales.

En una segunda etapa los leucocitos atraviesan el endotelio y migran hacia los lugares de infección o inflamación bajo la dirección de estímulos quimiotácticos. Los

quimioatrayentes, especialmente el sistema formado por las quimioquinas y los quimioatrayentes lipídicos junto con sus GPCRs (Fig.1.1) (*G protein-coupled receptors*) constituyen un grupo muy diverso. Desencadenan los procesos de adhesión intravascular de los leucocitos a las células del endotelio, y se encargan de guiarlos en su migración intersticial (Luster AD 2005).

Los patrones de migración son complejos y dependen no sólo del tipo de célula, sino también de su estado de diferenciación o de activación. Además, el endotelio vascular varía según la zona del organismo, por lo que tiene gran influencia en la migración celular. En general, la migración de los leucocitos a través de los endotelios depende de la carga electrostática de la superficie de las células que interaccionan, del flujo hemodinámico y de la expresión de moléculas de adhesión tanto en los leucocitos como en el endotelio. Por esto, la migración leucocitaria tiene lugar en pequeñas vénulas endoteliales donde la carga de la superficie es más baja, el flujo hemodinámico es reducido y donde se expresan selectivamente moléculas de adhesión (Roitt I 1998).

### **1.1.2. Adhesión transitoria**

En determinadas situaciones las células han de contactar unas con otras durante breves períodos de tiempo, dichos contactos generan señales intracelulares y se traducen en cambios fenotípicos en la célula. Algunos de estos procesos de comunicación celular transitorios tienen lugar durante la denominada respuesta inmunitaria específica, que incluye reacciones de adhesión celular entre células presentadoras de antígeno y linfocitos T, así como entre linfocitos B y T. Estos procesos de adhesión son esenciales para el desencadenamiento tanto de respuestas inmunes humorales, en las que los linfocitos B producen anticuerpos, como de respuestas celulares, que incluye tanto a linfocitos T colaboradores (Th) o citotóxicos (Tc) (Alberts B 1994).

Los procesos relacionados con la respuesta inmune humoral conducen a la generación de anticuerpos necesarios para la neutralización de virus y de las toxinas bacterianas, así como mediante el reclutamiento del sistema del complemento y de varios tipos de glóbulos blancos de la sangre que matan a los microorganismos extracelulares y a parásitos más grandes (Alberts B 1994). La producción de anticuerpos es la culminación de una serie de reacciones de adhesión celular que ocurren en el siguiente orden (Roitt I 1998):

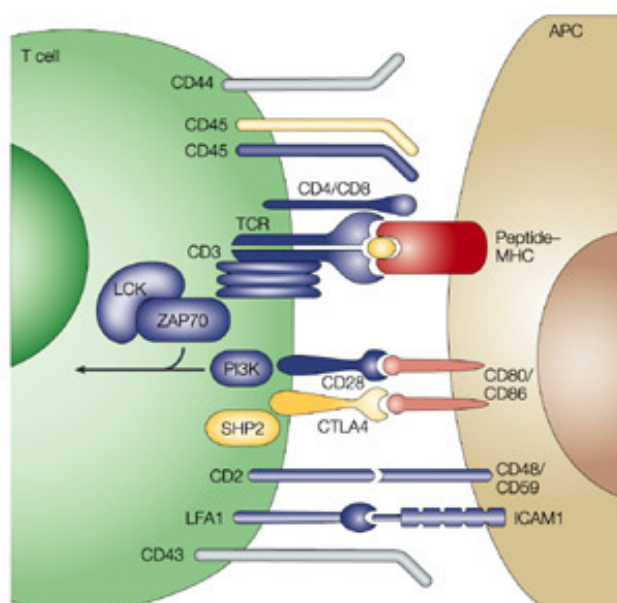


- Activación de las células T mediante el reconocimiento de antígeno presentado por las llamadas células presentadoras de antígeno (APC).
- Las células T colaboradoras interactúan con las células B.
- Las células B activadas proliferan y se diferencian en células productoras de anticuerpos.
- Los anticuerpos producidos inician distintas reacciones inmunitarias dependiendo del tipo que sean.

La respuesta inmune de tipo celular desempeña una importante función como mecanismo inmunológico de defensa, actuando principalmente contra bacterias y virus, así como evitando la aparición y desarrollo de células tumorales. En ella participan los linfocitos T colaboradores y citotóxicos. Los linfocitos reconocen el antígeno mediante el receptor de la célula T (TCR), una vez que dicho antígeno ha sido procesado en el interior de las células presentadoras de antígeno y expuesto en la superficie por moléculas del complejo principal de histocompatibilidad (MHC) (Garboczi DN 1996). Las moléculas MHC son una serie de glicoproteínas presentes en las membranas celulares, que se dividen en clase I y clase II. Los linfocitos Tc  $CD8^+$  reconocen a los antígenos que han sido procesados endógenamente en el citoplasma de la célula infectada y presentados en superficie por moléculas MHC de clase I, mientras que los linfocitos Th  $CD4^+$  interactúan con antígenos presentados por moléculas MHC de clase II. Cuando tiene lugar el reconocimiento antigénico entre el TCR y la molécula MHC que porta el antígeno, se desencadena una cascada de reacciones bioquímicas en el citoplasma de la célula T, conduciendo a la activación, proliferación y diferenciación celular. La consecuencia final de este tipo de respuesta es la formación de células T colaboradoras productoras de interleuquinas y de células T citotóxicas que poseen capacidad de lisar a las células que portan el antígeno que indujo su activación (Alberts B 1994).

Moléculas de adhesión celular como selectinas, integrinas e inmunoglobulinas son fundamentales para el desarrollo de estos procesos, manteniendo los contactos entre los diferentes tipos celulares en los procesos de presentación de antígeno (Fig.1.2). La organización del lugar de contacto entre las células juega un papel muy importante en el control de la inmunidad. El reconocimiento del antígeno por parte de la célula T requiere la formación de unas uniones especializadas conocidas como “sinapsis

inmunológicas” (Fig.1.2) (Grakoui A 1999), llamadas así por analogía con las sinapsis llevadas a cabo por las células del sistema nervioso. La sinapsis inmunológica está constituida por un grupo de receptores T que se unen a antígenos capturados por moléculas MHC, rodeados por un anillo de moléculas de adhesión en la superficie de la célula T que interaccionan con otras expresadas por la célula presentadora de antígeno (Fig.1.2) y su formación es el resultado de un mecanismo activo y dinámico que permite a las células T distinguir entre distintos ligandos antigénicos (Bretscher 1999; Grakoui A 1999). El estadio final de formación de una sinapsis inmunológica, es la



**Fig.1.2. Sinapsis inmunológica.** El reconocimiento del antígeno, necesario para la activación de la célula T, requiere la formación de unas uniones especializadas entre éstas y las células presentadoras de antígeno, conocidas como “sinapsis inmunológicas”. El complejo formado por TCR/MHC se sitúa en el centro de la interacción, rodeado por las interacciones entre las diferentes moléculas de adhesión formando un anillo. En la figura se indica la interacción TCR/MHC así como las diferentes moléculas que pueden formar parte de una sinapsis inmunológica típica.

estabilización de numerosos complejos de péptido-MHC con los receptores de la célula T y sus correspondientes moléculas de adhesión rodeándolos durante el tiempo necesario para la completa activación de la célula T (Fig.1.2).

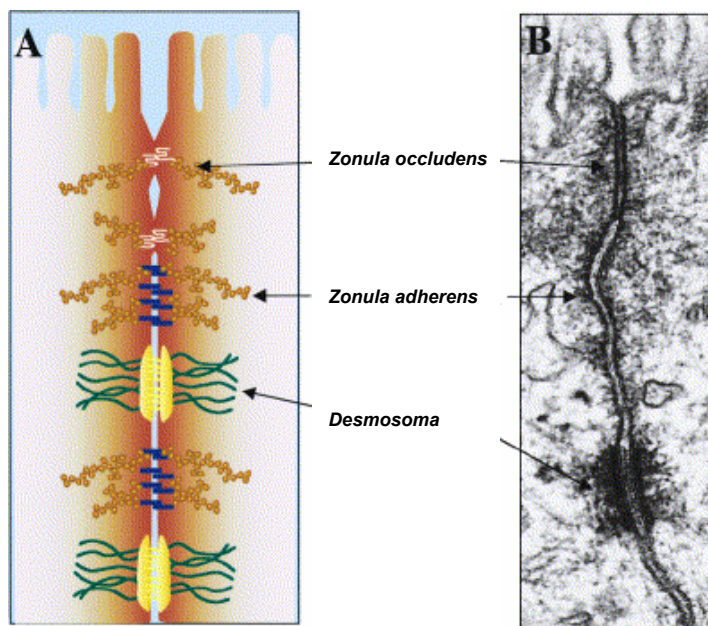
Un ejemplo de sinapsis inmunológica es el que llevan a cabo las células dendríticas en su interacción con los linfocitos T. Una vez que las células dendríticas llegan a los nódulos linfáticos, presentan sus antígenos en forma de péptidos unidos al MHC de tipo II a las células T inmaduras. El contacto inicial entre las células dendríticas y las células T parece establecerse mediante interacciones de DC-SIGN con ICAM-3 (Geijtenbeek TB 2000). Esto facilita la formación de uniones débiles entre moléculas de LFA-1 de las células T con moléculas de ICAM-1 de las células dendríticas. El complejo formado por TCR/MHC, junto con interacciones CD2/LFA-3, se sitúa en el centro de la “sinapsis inmunológica”, con las uniones LFA-1/ICAM-1 formando el anillo exterior de la misma (Bleijds DA 2001) (ver apartado 1.2.3 Fig.1.5B). Se piensa que la unión DC-SIGN-

ICAM-3 estabiliza el contacto de las membranas de las células T y dendríticas, aumentando la eficiencia del resto de receptores de la célula T. El hecho de que la unión de DC-SIGN-ICAM-3 sea de tipo transitorio permite a una célula dendrítica interactuar con un gran número de células T diferentes (Geijtenbeek TB 2000).

Se ha descrito un amplio espectro de células que pueden presentar antígenos, dependiendo de dónde y cómo tenga lugar el primer contacto entre el antígeno y las células del sistema inmune. Las células dendríticas, que se encuentran de forma abundante en los ganglios linfáticos y el bazo, son las células presentadoras de antígeno más efectivas en la activación inicial de las células T  $CD4^+$ . Sin embargo, macrófagos y células B también pueden llevar a cabo esta función.

### 1.1.3. Adhesión estática

La capacidad de las células para agruparse y formar tejidos viene proporcionada por interacciones de adhesión entre las mismas. Las interacciones de adhesión se localizan en regiones de contacto célula-célula y célula-matriz, que incluyen regiones de las membranas plasmáticas altamente especializadas.



**Fig.1.3. Diferentes tipos de uniones celulares.** A. Diagrama de los tres principales tipos de uniones en células epiteliales. Las uniones ocluyentes (*zonula occludens*) están compuestas por proteínas transmembrana unidas al citoesqueleto de actina y constituyen una barrera física entre las regiones apical y basolateral de las células. Las uniones de anclaje (*zonula adherens*) están formadas por interacciones homofílicas de cadherinas transmembrana unidas al citoesqueleto de actina. Los desmosomas se forman por interacciones puntuales de cadherinas unidas a los filamentos intermedios. B. Micrografía electrónica que muestra la ultraestructura de estos tres tipos de unión celular. (Figura adaptada de Pérez-Moreno, M. 2003) (Perez-Moreno M 2003).

Las uniones celulares se han caracterizado en detalle entre células epiteliales y se pueden dividir en tres tipos funcionalmente distintos:

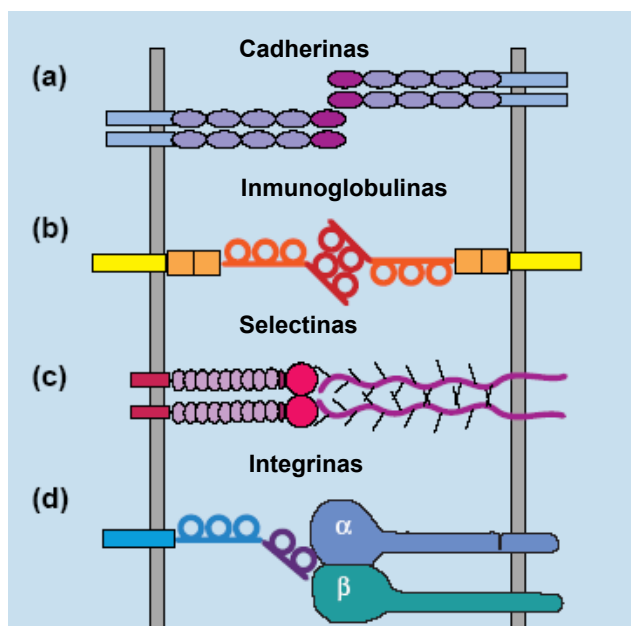
- **Uniones de oclusión (o estancas):** son uniones que sellan el espacio intercelular. Es la porción más apical del complejo de unión y se denomina *zonula occludens*, donde se establece una unión íntima entre las membranas plasmáticas de las células vecinas (Fig.1.3). Las hemimembranas externas de ambas células aparecen prácticamente en contacto. Tiene una profundidad aproximada de 1  $\mu\text{m}$  y forma un cinturón alrededor de la célula, lo que impide el paso de cualquier sustancia (Paniagua R 2003). A pesar de las diferencias existentes entre los diferentes tipos de epitelios todos ellos tienen la misión común de constituir barreras selectivas de permeabilidad. Un buen ejemplo son las uniones de oclusión entre las células epiteliales del intestino, que fuerzan el paso de las sustancias a través de la membrana de las microvellosidades (Alberts B 1994). Estas uniones son impermeables a la mayoría de moléculas, pero permiten el paso de algunos iones dependiendo del tipo celular. También evitan que receptores y proteínas de transporte ubicadas en la membrana se difundan a las superficies laterales o basales (Paniagua R 2003).
- **Uniones de anclaje:** unen mecánicamente las células (y sus citoesqueletos) a sus vecinas y a la matriz extracelular. Se dividen a su vez en dos tipos:
  1. **Unión adherente entre células:** recibe el nombre de *zonula adherens* y se encuentra inmediatamente a continuación de la *zonula occludens* en el complejo de unión. Se produce un ensanchamiento del espacio intercelular (aproximadamente 25 nm), que forma también un cinturón alrededor de la célula (Fig.1.3). Se forman por interacciones homofílicas entre moléculas de Cadherina-E transmembrana unidas al citoesqueleto de actina, también participan en la unión otras proteínas como Nectin-2 (Perez-Moreno M 2003).
  2. **Unión intercelular:** estas uniones se denominan *desmosomas* o *maculas adherens*, se sitúan a continuación de la *zonula adherens*. A diferencia de las anteriores no forma un cinturón alrededor de la célula sino una mancha o área de entre 0.5 y 1  $\mu\text{m}$  (Fig.1.3). Tiene como misión el anclaje de los filamentos intermedios y el aumento de la

cohesión entre células vecinas. Consta de dos partes, una externa adosada a la membrana, (placa), y un interna (zona fibrilar) (Paniagua R 2003). En las uniones participan glicoproteínas transmembrana del tipo cadherina unidas a los filamentos intermedios, las cuales emergen al espacio intercelular e interaccionan con las cadherinas de la célula adyacente (Paniagua R 2003).

- **Uniones de comunicación:** permiten el paso de señales químicas o eléctricas entre células adyacentes (Alberts B 1994). Se conocen con el nombre de *uniones de hendidura, nexo o acoplamiento*. En ellas, los 2 nm de separación están atravesados por numerosas partículas de contorno hexagonal que unen las membranas de las células adyacentes, cada una de ellas se denomina *conexón* y constituye una unidad funcional de la unión.

## 1.2. Moléculas de adhesión celular

Las moléculas de adhesión son proteínas presentes en la membrana plasmática que median procesos de interacción y comunicación celular (Fig.1.4). A menudo estas proteínas atraviesan la membrana y están sujetas al citoesqueleto. En algunos casos sirven como traductores de señal (Hynes 1999). Una misma molécula de adhesión



**Fig.1.4. Principales tipos de moléculas de adhesión.** a. **Cadherinas**, tienen cinco dominios extracelulares característicos, participan en adhesiones homofílicas dependientes de  $\text{Ca}^{2+}$ . b. **Inmunoglobulinas**, se caracterizan por poseer varios dominios inmunoglobulina concatenados. Familia muy diversa, con más de 100 tipos en vertebrados. c. **Selectinas**, juegan un papel fundamental en la adhesión de leucocitos al endotelio, como la interacción heterofílica mostrada entre una P-selectina y su receptor, la proteína glicosilada PSGL-1. La unión se realiza a través de un dominio lectina de tipo C (pintado en magenta). d. **Integrinas**, Están formadas por heterodímeros de dos polipéptidos ( $\alpha$  y  $\beta$ ) transmembrana. Se agrupan en tres subfamilias principales, dependiendo del tipo de sus cadenas  $\beta$  ( $\beta_1$ ,  $\beta_2$  ó  $\beta_3$ ). Se muestra una interacción heterofílica entre el receptor ICAM-1 y la integrina  $\alpha_1\beta_2$  (LFA-1). (Figura adaptada de Hynes, R.O. 1999) (Hynes 1999).

puede unirse a más de un ligando mediante sitios de unión diferentes. Aunque la afinidad de las moléculas individuales por los ligandos suele ser baja, los procesos de adhesión celular implican uniones multivalentes y por lo tanto de alta avidéz.

Las interacciones de adhesión celular pueden ser moduladas ya sea por cambios en el número de moléculas de adhesión en la superficie celular o por cambios en su afinidad por el ligando. El nivel de expresión de moléculas de adhesión en superficie se puede cambiar mediante dos mecanismos: La transferencia rápida de una gran cantidad de moléculas de adhesión retenidas en vesículas, o bien por la síntesis de nuevas moléculas, proceso mucho más lento y que puede llevar varias horas. Cambios conformacionales en las moléculas de adhesión celular modifican su afinidad por los ligandos, tal y como se ha descrito con las integrinas (Hynes 1992; Lee JO 1995; Qu A 1995).

Aunque han sido identificadas un gran número de moléculas de adhesión, la mayoría pueden clasificarse en cuatro familias estructuralmente distintas: cadherinas, selectinas, integrinas e inmunoglobulinas (Fig.1.4) (Hynes 1999).

- **Cadherinas:** se han descrito aproximadamente 20 tipos de “cadherinas clásicas”, constan de cinco dominios extracelulares característicos (Fig.1.4A), cada uno de los cuales contiene un *sandwich* de láminas- $\beta$ . Las cadherinas median adhesiones homofílicas dependientes de  $\text{Ca}^{+}$ , a través de los dominios más distales. Comparten dominios citoplasmáticos homólogos que se unen al citoesqueleto de actina.
- **Inmunoglobulinas:** la superfamilia de las inmunoglobulinas (Ig) incluye un elevado número de receptores celulares implicados en diversos procesos celulares, muchos de ellos relacionados con el funcionamiento del sistema inmune. Estos receptores tienen uno o varios dominios Ig concatenados en su región extracelular (Fig.1.4B), dominios constituidos por un *sandwich* de dos hojas  $\beta$  comúnmente unidas entre si por un puentes disulfuro. A pesar de la semejanza estructural entre los receptores Ig, la similaridad en secuencia es realmente baja (<20%). Dentro de esta extensa familia se han definido grupos de receptores o subfamilias que desempeñan funciones similares y/o presentan una mayor similaridad en secuencia. Uno de estos grupos comprende moléculas que se unen a integrinas y que están implicadas en procesos de adhesión intercelulares dinámicos, como la unión de leucocitos a

células endoteliales en el torrente sanguíneo. En este grupo nos encontramos los ICAMs (*intercellular adhesion molecules*), VCAM-1 (*vascular cell adhesion molecule-1*) y MAdCAM-1 (*mucosal addressin cell adhesion molecule-1*). MAdCAM-1 es una molécula peculiar ya que contiene dos Ig seguidos por un dominio de mucina glicosilado en la región extracelular, lo que le permite unirse tanto a integrinas como a selectinas.

- **Integrinas:** es un grupo amplio de moléculas de adhesión que están presentes en muy diversos tipos celulares, incluidos leucocitos. Al contrario de los otros tipos descritos están formadas por heterodímeros (Fig.1.4D). Constan de dos polipéptidos ( $\alpha$  y  $\beta$ ). Existen tres subfamilias principales dependiendo del tipo de sus cadenas  $\beta$  ( $\beta_1$ ,  $\beta_2$  ó  $\beta_3$ ), aunque el tipo de asociación  $\alpha\beta$  puede ser muy diverso. Las integrinas  $\beta_1$  participan en la unión de células a la matriz extracelular, las  $\beta_2$  están implicadas en la adhesión de leucocitos al endotelio o a otras células del sistema inmune, y las  $\beta_3$  median procesos de adhesión de plaquetas y neutrófilos en los sitios de inflamación o de daño vascular. Integrinas del tipo  $\beta_1$  y  $\beta_2$  tienen un papel importante en adhesión celular heterotípica, especialmente en leucocitos, donde se unen a receptores de la superfamilia de las inmunoglobulinas (ICAMs, VCAM-1, MAdCAM-1). Las integrinas que se unen a los ICAMs presentan un dominio de unión característico conocido como dominio insertado (I), mientras que las que se unen a VCAM-1 o MAdCAM-1 carecen de este dominio (Larson RS 1989). La unión de las integrinas a sus ligandos es dependiente de cationes divalentes.
- **Selectinas:** incluyen las moléculas selectina-E y selectina-P, que se expresan en endotelio y plaquetas respectivamente, y la selectina-L, que se expresa en algunos leucocitos. Son moléculas con una región transmembrana y con dominios extracelulares de diferentes tipos (Fig.1.4C). El dominio N-terminal es del tipo lectina de tipo C (Hynes 1999), por lo que las selectinas se unen a ligandos que presentan carbohidratos.

### 1.2.1. La subfamilia de los ICAMs

La subfamilia de los ICAMs comprende un grupo de 6 receptores celulares que reconocen la integrina leucocitaria LFA-1 ( $\alpha_L\beta_2$ ). LFA-1, (*Lymphocyte Function-Associated Antigen-1*), es una proteína integral de membrana perteneciente a familia de las integrinas (Larson RS 1989). Participa en numerosos procesos dentro del sistema inmune, como en interacciones célula-célula durante la migración de los linfocitos a través de los endotelios hacia los lugares de inflamación y durante la recircularización a través de los nódulos linfáticos (Patarroyo M 1990; Springer 1990; Springer 1994). También está presente en otros procesos como: la activación de linfocitos T, la adhesión intercelular de leucocitos o la unión de linfocitos T citotóxicos a las células presentadoras de antígeno. Está constituida por una subunidad  $\alpha$  (CD11a) de entre 120 y 180 KDa y una subunidad  $\beta$  (CD18) con un tamaño comprendido entre 90 y 110 KDa, asociadas de manera no covalente. Posee el característico dominio “I” (insertado) (Larson RS 1989) con un sitio de adhesión MIDAS (*metal ion-dependent adhesion site*) (Lee JO 1995). Se expresa en todos los leucocitos de forma constitutiva, pero necesita ser activada para unirse a sus ligandos. La activación se lleva a cabo mediante señalización intracelular y la influencia de cationes.

Los ICAMs son proteínas de membrana del tipo I con un único segmento transmembrana y con una región extracelular formada por dominios inmunoglobulina (Ig) concatenados (Fig.1.4B). Hay una gran diferencia en el número de dominios inmunoglobulina entre los miembros de la subfamilia, que van desde dos en ICAM-2 e ICAM-4 a nueve en ICAM-5 (Fig.1.6). Comparten homologías estructurales y funcionales y son importantes en interacciones célula-célula fundamentales para el funcionamiento del sistema inmune (Springer 1990). Todos ICAMs excepto ICAM-4 tienen un ácido glutámico en el dominio N-terminal que es esencial para la unión a LFA-1.

Las interacciones de adhesión celular mediadas por miembros de la familia de los ICAMs y las integrinas leucocitarias LFA-1 y Mac-1 regulan el tráfico de los leucocitos entre tejidos inflamados y no inflamados y contribuyen a la respuesta antigénica específica de las células T (Fig.1.5). Recientemente se ha descrito la unión de ICAM-2 e ICAM-3 a la lectina DC-SIGN (*dendritic cell-specific ICAM-grabbing nonintegrin*), unión importante para el funcionamiento de las células dendríticas (Berg JM 2002) (Fig.1.5).

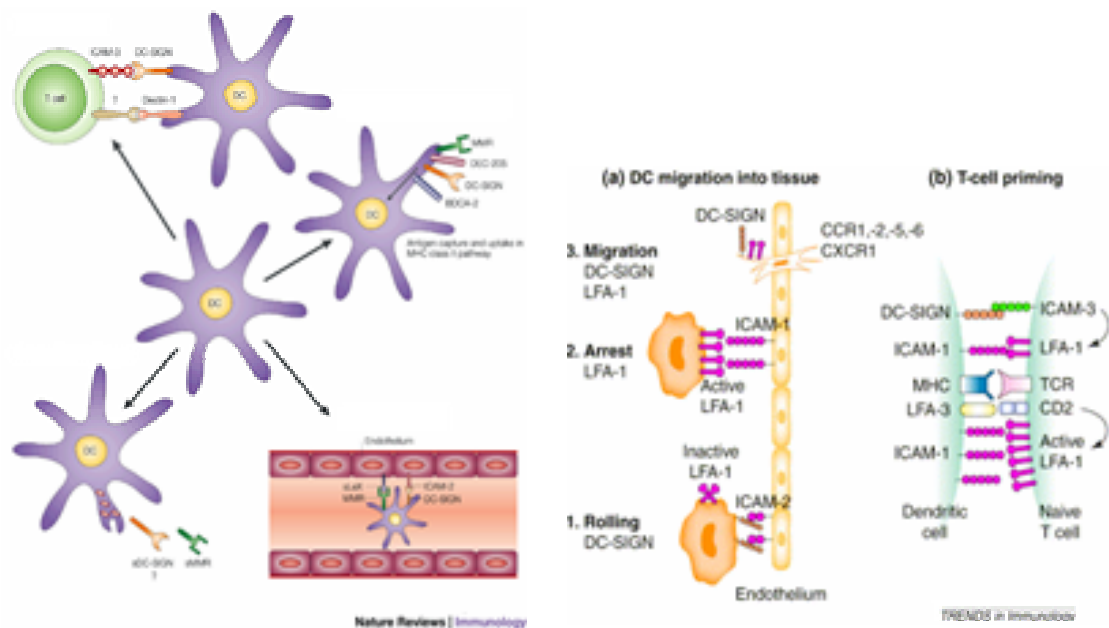


### 1.2.2. ICAM-1

ICAM-1 es un receptor cuya expresión es inducible por citoquinas y que se une a las integrinas leucocitarias LFA-1 y Mac-1. Se distribuye por una gran variedad de tejidos y la presencia de citoquinas inflamatorias activa su expresión en leucocitos, células epiteliales y endoteliales (Bleijns DA 2001). La principal función de ICAM-1 es proporcionar adhesión entre células endoteliales y leucocitos tras una lesión o en situaciones de estrés, estando asociado a procesos inflamatorios (Bella J 1999). Asimismo ICAM-1 es usado como receptor por el principal grupo de rinovirus humano (HRV) y es fundamental para la entrada del virus en la célula (Greve JM 1989; Staunton DE 1989). Contiene cinco dominios Ig en la parte extracelular (Geijtenbeek TB 2000). Se han descrito las estructuras tridimensionales de fragmentos de ICAM-1 con los dos dominios N-terminales y con los dominios 3 a 5, mientras que la molécula completa presenta cierta flexibilidad entre los dominios 2 y 3 (Casasnovas JM 1998; Yang Y 2004). El primer dominio de ICAM-1 es reconocido por la integrina LFA-1 así como por otros ligandos tales como HRV o proteínas derivadas de *Plasmodium falciparum*. Este dominio carece de glicosilaciones (Fig.1.7B), en contraste con el resto de los ICAMs humanos. Recientemente se publicó la estructura cristalográfica de un complejo del dominio I de LFA-1 y ICAM-1, que define en detalle la interacción de adhesión celular. Un catión de  $Mg^{2+}$  que se encuentra en el MIDAS del dominio I de LFA-1 aparece coordinado con el ácido glutámico conservado en el dominio N-terminal de los ICAMs, residuo esencial para la unión a la integrina (Lee JO 1995).

### 1.2.3. ICAM-2

ICAM-2 es un receptor que se expresa constitutivamente en endotelio, de forma que puede ser usado para la recircularización basal de leucocitos, no estando asociado a inflamación. Contiene solamente dos dominios Ig en la parte extracelular, para los que se conoce su estructura (Casasnovas JM 1997). La región extracelular de ICAM-2 está altamente glicosilada, lo que posiblemente determine su capacidad de unión a la lectina DC-SIGN, objeto de estudio en este trabajo. La unión de DC-SIGN a ICAM-2 o ICAM-3 requiere la presencia calcio, característica de las uniones de lectinas de tipo C (Geijtenbeek TB 2000; Jiménez D 2005). Esta interacción parece ser relevante para la biología de las células dendríticas.



**Fig.1.5.** Las células dendríticas son células especializadas en el reconocimiento de patógenos y tienen un papel fundamental en el control de la inmunidad. La lectina DC-SIGN, expresada por las células dendríticas participa mediante su unión a ICAM-2 e ICAM-3 en los procesos de trasmigración y activación de la célula T, respectivamente. **A. Migración de las células dendríticas.** 1. Las células dendríticas “ruedan” por el endotelio gracias a las interacciones ICAM-2/DC-SIGN. 2. Posteriormente se fijan a las paredes endoteliales mediante interacciones ICAM-1/LFA-1. 3. En un proceso mediado por quimioquinas las células dendríticas son capaces de migrar a través de los endotelios de forma específica. **B.** Durante el proceso de migración las células dendríticas maduran y llegan a los nódulos linfáticos, donde se adhieren a las células T inmaduras mediante interacciones ICAM-3/DC-SIGN. Estas interacciones facilitan uniones ICAM-1/LFA-1, MHC/TCR y LFA-2/CD2, estabilizando la unión conocida como “sinapsis inmunológica”. (Figura adaptada de Figdor CG, 2002 y Bleijs DA, 2001) [2, 3].

Las células dendríticas actúan como guardianes del sistema inmune, procesando y presentando antígenos a los linfocitos. La maduración de las células dendríticas se lleva a cabo mediante la captura de antígeno, así como por la unión a citoquinas y quimioquinas inflamatorias, lo cual induce la maduración y migración de las células desde la periferia a los órganos linfáticos, donde entran en contacto con células T específicas (Fig.1.5) (Figdor CG 2002). No se sabe mucho sobre las bases moleculares de la migración de las células dendríticas y se piensa que las elevadas cantidades de DC-SIGN presentes en su superficie podrían estar implicadas en estos procesos.

La trasmigración a través de los endotelios sólo puede darse si las células son capaces de unirse a las paredes de los vasos endoteliales, venciendo la fuerza tangencial del flujo sanguíneo (Springer 1994). Para ello las células en la sangre han de disminuir su velocidad, interaccionando y “rodando” por las paredes del endotelio. Como ya hemos descrito, en este proceso intervienen las selectinas presentes en leucocitos o en

las paredes endoteliales que reconocen carbohidratos (Vestweber D 1999; Chang KC 2000). Además de la posible intervención de las selectinas en el proceso, las células dendríticas pueden unirse y “rodar” por el endotelio mediante las interacciones de DC-SIGN con ICAM-2 (Fig.1.5A). La subsiguiente activación de la integrina LFA-1, que se une a moléculas de ICAM-1 o ICAM-2 presentes en las paredes del endotelio, mediará el proceso de trasmigración de la célula dendrítica (Bleijs DA 2001).

#### **1.2.4. ICAM-3.**

Al igual que ICAM-1, con el que guarda una elevada similaridad, ICAM-3 tiene 5 dominios Ig en la región extracelular. No obstante, a diferencia de ICAM-1, ICAM-3 se expresa en leucocitos y no en endotelio. Las integrinas LFA-1 ( $\alpha_L\beta_2$ ) y  $\alpha_D\beta_2$  han sido descritas como los receptores primarios de ICAM-3, aunque su afinidad por ambas integrinas, incluso después de la activación de las mismas, es relativamente baja (Dustin ML 1989). Recientemente se ha determinado la estructura de un complejo del dominio N-terminal de ICAM-3 con el dominio I de LFA-1 (Song G 2005). La interacción de ICAM-3 con LFA-1 está relacionada con el inicio de la respuesta inmune.

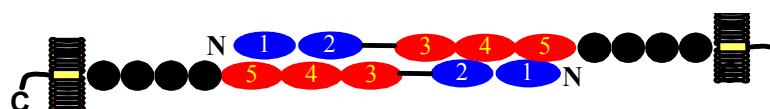
Estudios recientes han mostrado la unión de ICAM-3 a la lectina DC-SIGN, interacción que media procesos de adhesión entre células dendríticas y células T y que conduce a la activación de células T (Fig.1.5B) (Geijtenbeek TB 2000). La unión de ICAM-3 a DC-SIGN es dependiente de  $\text{Ca}^{2+}$  y puede ser inhibida por manosa (Geijtenbeek TB 2000; Jiménez D 2005), lo que es característico de las lectinas de tipo C, que poseen dominios de reconocimiento de carbohidratos extracelulares (CRDs). La elevada expresión de ICAM-3 en linfocitos T inmaduros sugiere que puede mediar el contacto inicial entre las células dendríticas y las células T, iniciando su activación y la formación de la “sinapsis inmunológica” por el complejo TCR/MHC y numerosas interacciones de adhesión circundantes (Fig.1.5B) (Bleijs DA 2001).

#### **1.2.5. ICAM-5**

ICAM-5, llamado inicialmente telencefalina, es una glicoproteína integral de membrana del tipo I de 130-kDa, que se expresa exclusivamente en neuronas telencefálicas de mamíferos (Mizuno T 1997). El telencéfalo es el lugar donde reside la

coordinación motora, el conocimiento superior y la consciencia en los humanos, siendo la región más compleja del cerebro adulto de los mamíferos (Guillemot 2005). Durante el desarrollo embrionario el telencéfalo se expande de forma muy notable en los mamíferos y forma las complejas estructuras telencefálicas del cerebro adulto (Mori K 1987). ICAM-5 participa en el desarrollo cerebral y la sinaptogénesis ya que su expresión se inicia con el nacimiento, en paralelo con la maduración del sistema nervioso central y el desarrollo de los complejos circuitos neuronales. En particular se ha visto que ICAM-5 promueve el desarrollo de neuritas (Tian L 2000), lo que indica un papel fundamental en el desarrollo del sistema nervioso. Asimismo, hay estudios que prueban el aumento de ICAM-5 soluble en suero de pacientes con encefalitis aguda así como la reducción de la actividad de ICAM-5 en cerebros de pacientes enfermos de Alzheimer, relacionando esta molécula con patologías cerebrales (Tian L 2000).

ICAM-5 tiene nueve dominios inmunoglobulina con un total de 832 aminoácidos en la región extracelular, un segmento transmembrana de 28 aminoácidos y un dominio citoplásmico de 64 aminoácidos (Mizuno T 1997). Es muy homólogo a los otros miembros de la subfamilia de los ICAMs, con los que comparte una identidad de entre 38-55% (Tian L 2000). Como otros miembros de la subfamilia se une a la integrina leucocitaria LFA-1, lo cual indica que puede ser importante en reacciones inmunitarias localizadas en el sistema nervioso central (Tian L 2000). Se ha descrito la adhesión entre células T y neuronas mediada por la unión de ICAM-5 a LFA-1 y que en esta



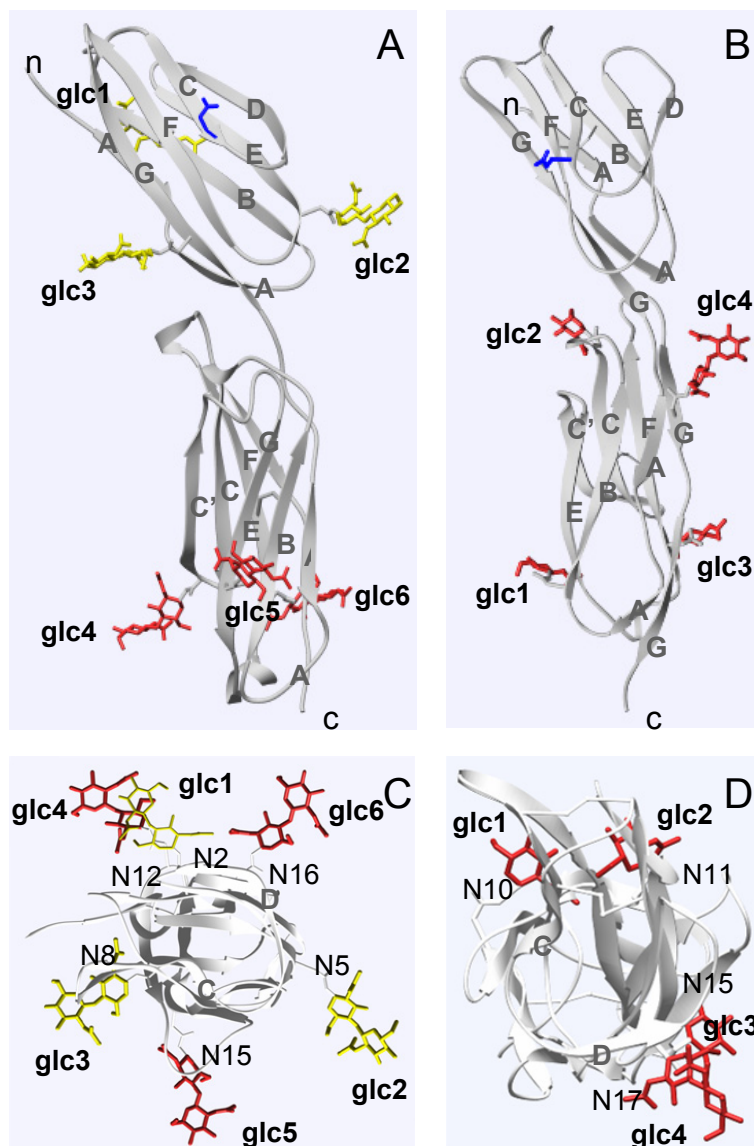
**Fig.1.6. Unión homofílica de IC5.** Los dominios básicos 1 y 2 (pI 11) indicados en azul, interaccionan de forma homofílica con los dominios ácidos 3,4 y 5 (pI 4) pintados en rojo.

unión el dominio 1 es fundamental (Tian L 2000). Aparte de estas funciones compartidas con otros ICAMs, se ha descrito una interacción de adhesión homofílica exclusiva de ICAM-5. La interacción entre moléculas de ICAM-5 localizadas en neuronas opuestas promueve el desarrollo de neuritas, relacionado con el proceso de sinaptogénesis. Estas interacciones de adhesión homofílicas implican a los 5 dominios N-terminales del receptor, que presentan singularidad en su carga (Tian L 2000). Los dos dominios N-terminales (1 y 2) son altamente básicos (pI 11), mientras los dominios 3,4 y 5 son ácidos (pI 4). Así pues, se ha propuesto que la interacción de adhesión está

fundamentada en una interacción electrostática entre los módulos ácido y básico de ICAM-5 (Fig.1.6).

### 1.3. N-glicosilaciones en los miembros de la subfamilia de los ICAMs

La mayoría de los dominios inmunoglobulina de los ICAMs poseen glicosilaciones de tipo N y algunos de ellos están altamente glicosilados (Fig.1.7). Es posible que estos azúcares participen del plegamiento y reconocimiento de ligandos, objeto de análisis en esta tesis. La supresión del primer sitio de glicosilación en ICAM-3 afecta significativamente la conformación de varios epítomos de unión a anticuerpos así como la unión del receptor a LFA-1, mientras que la eliminación de otros azúcares del primer dominio no tiene efectos en la unión de ligandos (Klickstein LB 1996). La mutación de varios N-glicanos en el segundo dominio de ICAM-1 afectó marginalmente la unión con LFA-1 (Staunton DE 1990). La resolución de las estructuras tridimensionales de los dos dominios N-terminal glicosilados de ICAM-1 e ICAM-2 (Fig.1.7) mostraron la localización de los azúcares en los dominios y su conformación (Casasnovas JM 1997; Casasnovas JM 1998). Tanto el primero como el segundo residuo de GlcNAc del núcleo del azúcar estaban bien definidos en los mapas de densidad electrónica. Sin embargo no se pudo observar la densidad correspondiente al tercer residuo de manosa, lo que sugiere que estos residuos son muy flexibles. La distribución de los azúcares es diferente en ICAM-1 e ICAM-2 (Fig.1.7). En ICAM-2 los azúcares se distribuyen homogéneamente alrededor de los dos dominios (Fig.1.7A, C), mientras que en ICAM-1 se localizan en dos caras opuestas del segundo dominio (Fig.1.7B, D). Los azúcares se localizan en la parte inferior de los dominios Ig de ICAM-2 y tienen una distribución tipo trípode en el segundo dominio, lo que sugiere que pueden servir para orientar a la molécula en la superficie celular (Casasnovas JM 1997). En ICAM-1 el primer dominio carece de glicosilaciones (Fig.1.7B), exponiendo una mayor superficie proteica para la dimerización en *cis* y facilitando uniones a diversos ligandos como HRV y proteínas derivadas de *plasmodium falciparum* (Bella J 1998; Casasnovas JM 1998).



**Fig.1.7. Estructuras de IC1 e IC2.** Modelo con los dos dominios Ig de ICAM-2 (A y C) y de ICAM-1 (B y D) coloreados en gris, las N-glicosilaciones del primer y segundo dominio están pintadas en amarillo y rojo, respectivamente. Se ha visto que moléculas solubles de IC1 con los dos dominios inmunoglobulina N-terminal son funcionales en la unión a ligandos. El primer dominio de IC1 carece de glicosilaciones, a diferencia del resto de los ICAMs humanos.

#### 1.4. Estructura, biosíntesis y función de los carbohidratos.

Los grupos carbohidratos se unen de forma covalente a un gran número de proteínas conocidas como glicoproteínas. Los azúcares se unen al átomo de nitrógeno de la cadena lateral de una asparagina, (unión de tipo N), o al átomo de oxígeno de la cadena lateral de una serina o una treonina (unión de tipo O). Un residuo de asparagina puede aceptar un oligosacárido si forma parte de una secuencia de aminoácidos formada por

Asn-X-Ser o Asn-X-Thr, donde la X puede ser cualquier residuo. Por tanto, los posibles sitios de glicosilación se pueden predecir a partir de la secuencia de aminoácidos de la proteína. Aunque ello no significa que todos estos sitios estén glicosilados, ya que esta modificación depende de la estructura de la proteína y del tipo celular donde se exprese. Todos los oligosacáridos de tipo N tienen en común un núcleo central formado por un pentasacárido consistente en tres manosas y dos N-acetilglucosaminas. A este núcleo se pueden unir azúcares adicionales formando una gran variedad de oligosacáridos diferentes (Berg JM 2002).

Los azúcares de tipo N presentes en glicoproteínas pueden ser de tipo complejo, híbrido o de alto contenido en manosas (Ashford DA 1993; Dwek 1995; Rudd PM 2001). Los azúcares con alto contenido en manosas son precursores biosintéticos de los otros dos. Todos comparten el pentasacárido  $\text{Man}_3\text{GlcNAc}_2$ . Por el contrario, el único elemento común de los azúcares de tipo O es la N-acetilgalactosamina unida al grupo hidroxilo de la serina o la treonina. En ambos tipos de cadenas oligosacarídicas, la extensión de las mismas se suele producir por adición de la secuencia  $(\text{Gal}\beta 1,4)_3\text{GlcNAc}\beta 1,3$  repetida, seguida de varios azúcares terminales. La mayor diversidad de las cadenas oligosacarídicas se produce por la variación de los azúcares terminales, que pueden ser residuos de ácido siálico, fucosa, galactosa, N-acetilglucosamina o N-acetilgalactosamina, solos o en combinación con uno o más grupos hidroxilo (Paulson 1989). El número de combinaciones es muy elevado, por lo que los glucanos tienen una gran diversidad estructural. Se sabe que la secuencia de los carbohidratos cambia dependiendo de los tejidos, e incluso de los tipos celulares dentro de un mismo tejido (Paulson 1989).

La glicosilación de proteínas de tipo N comienza en el retículo endoplásmico y finaliza en el complejo de Golgi, mientras que la de tipo O tiene lugar exclusivamente en el complejo de Golgi. Las glicoproteínas de tipo N adquieren sus azúcares a partir de una molécula lipídica compuesta por 20 unidades de isopreno ( $\text{C}_5$ ) llamado dolicol. Un oligosacárido destinado a unirse a una asparagina se une a una molécula de dolicol fosfato. El grupo fosfato terminal es el lugar donde se une el oligosacárido activado, que será transferido al aceptor proteico. El dolicol fosfato reside en la membrana del retículo endoplásmico con el fosfato orientado hacia el citoplasma. El proceso de ensamblaje ocurre en tres etapas. Inicialmente dos residuos de N-acetilglucosamina y cinco de manosa se unen al dolicol fosfato por medio de una serie de enzimas citoplasmáticas. Posteriormente esta estructura es trasladada a través de la membrana hacia el lumen del

retículo endoplásmico. Finalmente se añaden otra serie de azúcares usando monosacáridos activados unidos a dolicol fosfato. Este proceso termina con la formación de un oligosacárido de 14 residuos unido a dolicol fosfato. Este oligosacárido precursor es transferido en bloque a un residuo específico de asparagina de una cadena polipeptídica en crecimiento. Antes de abandonar el retículo endoplásmico, se eliminan tres moléculas de glucosa del oligosacárido de 14 residuos, lo que constituye un control de calidad que asegura que solo las moléculas correctamente plegadas continúen su procesamiento (Berg JM 2002).

Las proteínas son entonces transportadas al complejo de Golgi, donde los carbohidratos de las glicoproteínas son modificados de formas muy diversas. Las proteínas pasarán desde el complejo de Golgi hacia los lisosomas, los gránulos de secreción o la membrana plasmática de acuerdo con una serie de señales codificadas por su secuencia de aminoácidos y su estructura tridimensional. Los carbohidratos de tipo N de las glicoproteínas son modificados en cada uno de los compartimentos del complejo de Golgi, añadiendo o quitando residuos según el tipo de azúcar que se vaya a formar. La secuencia final del oligosacárido depende de las glucosiltransferasas presentes en el compartimiento del Golgi donde sean procesadas. Hay que destacar, que a pesar de todo el mecanismo de procesamiento, el pentasacárido inicial permanece invariable (Berg JM 2002).

Está generalmente aceptado que los azúcares previenen interacciones no específicas entre proteínas y protegen las glicoproteínas de proteasas, aumentando su estabilidad (Rudd PM 2001). Otro papel importante de los carbohidratos es el que juegan en el reconocimiento celular, en procesos de interacciones patógeno-huésped o interacciones célula-célula (Paulson 1989). Virus como el VIH utilizan las glicosilaciones para enmascarar epítomos proteicos y evitar la neutralización por anticuerpos del huésped (Kwong PD 2002).

### **1.5. Lectinas**

Las lectinas son proteínas que se unen específicamente a estructuras glucídicas. Se encuentran presentes en plantas, animales y microorganismos. La principal función de las lectinas en eucariotas superiores es facilitar el contacto entre células. Una lectina suele contener dos o más sitios de unión a carbohidratos (Fig.1.8B, C), aunque algunas



forman estructuras oligoméricas con múltiples sitios de unión. Las lectinas presentes en la superficie de una célula interactúan con carbohidratos presentes en la superficie de otra célula. Las interacciones son relativamente débiles, lo que sugiere que deben de mediar procesos de adhesión dinámicos (Berg JM 2002).

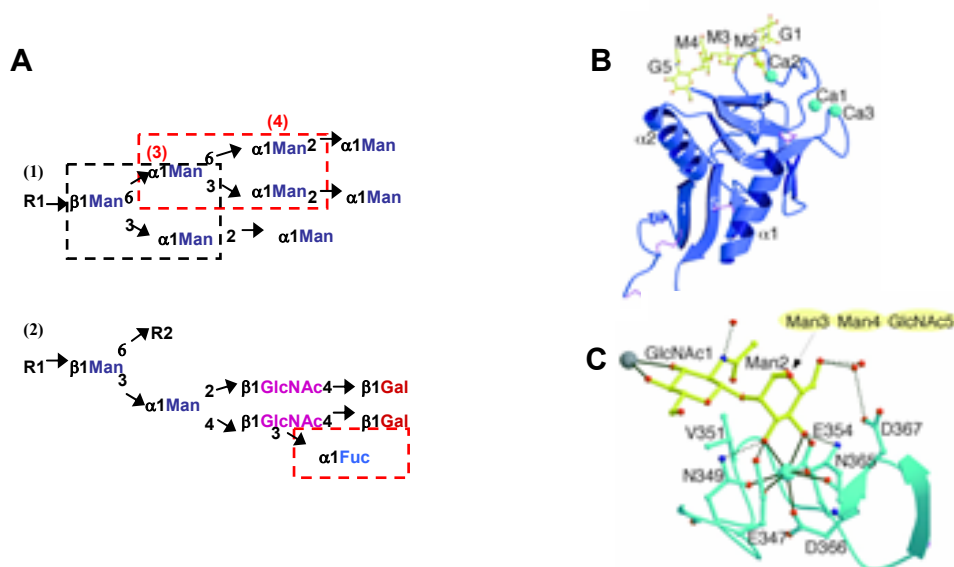
### 1.5.1. Lectinas de tipo C y sus ligandos

Las lectinas se pueden clasificar en función de su secuencia de aminoácidos y de sus propiedades bioquímicas (Berg JM 2002). Un tipo de lectina de interés para este trabajo es la del tipo C. Las lectinas de tipo C unen azúcares mediante un mecanismo dependiente de calcio, usando unos dominios de reconocimiento de carbohidratos (CRD) altamente conservados (Fig.1.8B). Estos dominios poseen unos compartimientos de unión a calcio fundamentales para el reconocimiento de carbohidratos (Figdor CG 2002). Un ion de calcio actúa como puente entre la proteína y el azúcar, coordinándose con grupos hidroxilo del carbohidrato y con residuos ácidos y polares del dominio CRD (Fig.1.8C). Residuos básicos de la proteína pueden formar puentes de hidrógeno con grupos hidroxilo del carbohidrato (Berg JM 2002).

Las lectinas de tipo C se pueden expresar como proteínas de membrana o secretarse de forma soluble. Como ejemplos de lectinas solubles son las proteínas pulmonares SP-A y SP-D, que se denominan colectinas y se secretan en la superficie luminal de las células epiteliales pulmonares. También se incluyen en este grupo las proteínas de unión a manosa (MBP). Se han caracterizado un gran número de lectinas de tipo C unidas a membrana, incluidas algunas que sólo se expresan en células dendríticas. En las células dendríticas se pueden distinguir dos grupos de lectinas de tipo C unidas a membrana, dependiendo de la orientación de su extremo N terminal. Son de tipo I o tipo II en función de que el N terminal esté colocado en el exterior o el citoplasma de la célula, respectivamente. Las lectinas de tipo I contienen varios dominios de unión a carbohidratos, mientras que las de tipo II tienen sólo uno (Figdor CG 2002).

### 1.5.2. DC-SIGN

DC-SIGN (*dendritic-cell-specific ICAM-3 grabbing non-integrin*) es una lectina de tipo C presente en la superficie de las células dendríticas. DC-SIGN se expresa en



**Fig.1.8. A .** Estructuras oligosacáridicas típicas que reconoce DC-SIGN. **(1)** Rica en manosas, en el recuadro rojo se muestra la estructura  $\text{Man}\alpha 1-3$  [ $\text{Man}\alpha 1-6$ ]  $\text{Man}$  reconocida por la lectina, también se indican en rojo los residuos de manosa equivalentes con los del oligosacárido de las figuras B y C. El recuadro negro señala la estructura  $\text{Man}\alpha 1-3$  [ $\text{Man}\alpha 1-6$ ]  $\text{Man}$  interna que no se une a la lectina debido a la conformación del azúcar. **(2)** Oligosacárido que contiene fucosa terminal, el recuadro rojo indica la zona de unión a la lectina. **B.** Estructura del dominio de unión a carbohidratos (CRD) de DC-SIGN unido a un pentasacárido  $\text{Man}_3\text{GlcNAc}_2$ . **C.** Detalle de la interacción de un pentasacárido sintético con DC-SIGN.

células dendríticas inmaduras en tejidos periféricos, así como en células maduras de los tejidos linfáticos, (nódulos linfáticos, amígdalas y bazo). Se ha detectado de forma abundante en células dendríticas presentes en la lámina propia de tejidos mucosos, como recto, útero... DC-SIGN juega un papel importante en el proceso de captura y presentación de antígeno (Geijtenbeek TB 2000).

Es una proteína de membrana de tipo II, cuyo dominio extracelular está formado por 23 residuos en tandem repetidos siete veces de forma completa y una de forma incompleta, así como de un CRD,  $\text{Ca}^{2+}$  dependiente, en el extremo COOH terminal (Figdor CG 2002). La región repetitiva constituye un tallo que media tetramerización (Feinberg H 2001). El dominio CRD de DC-SIGN adopta una conformación típica de las lectinas de tipo C (Drickamer 1999). La segunda  $\alpha$ -hélice es inusualmente larga y está más cerca del resto del dominio que en la mayoría de lectinas de tipo C. El extremo C-terminal se empaqueta contra el *loop* que conecta las láminas- $\beta$  3 y 4, formando una superficie continua que interacciona con el oligosacárido (Fig.1.8B) (Feinberg H 2001).

Las interacciones de DC-SIGN con glicanos endógenos se producen mediante la interacción con un motivo característico de carbohidratos ricos en manosas o con fucosilación terminal (Guo Y 2004) (Fig.1.8 A).

El mecanismo de reconocimiento de carbohidratos por DC-SIGN se describió por la estructura cristalográfica del dominio CRD unido a un pentasacárido sintético compuesto por tres manosas y dos N-acetilglucosaminas (G1M2M3M4G5) (Fig.1.8B y C) (Feinberg H 2001). La conformación del pentasacárido es la misma que la de los azúcares de alto contenido en manosas, cuyo núcleo central es del tipo  $\text{Man}\alpha 1-3$  [ $\text{Man}\alpha 1-6$ ] Man (Fig.1.8A1). El residuo manosa unido en conformación  $\alpha 1-3$ , Man2, (Fig.1.8B y C), se une al sitio  $\text{Ca}^{2+}$  presente en todas las lectinas de tipo C, los hidroxilos ecuatoriales 3- y 4- forman enlaces de coordinación con el  $\text{Ca}^{2+}$  y puentes de hidrógeno con los aminoácidos de la proteína (Fig.1.8C). Los dos residuos de manosa que forman la unión  $\alpha 1-6$ , Man3 y Man4 interaccionan mediante fuerzas de Van der Waals y puentes de hidrógeno con los residuos  $\text{Phe}^{325}$  y  $\text{Ser}^{372}$ , respectivamente. Los residuos de N-acetilglucosamina terminales también forman diferentes contactos con DC-SIGN. El primero se une al residuo  $\text{Val}^{351}$ , además interacciona con el sitio  $\text{Ca}^{2+}$  de una molécula de DC-SIGN adyacente, uniendo ambos monómeros (Fig.1.8C) (Feinberg H 2001).

Se ha demostrado que DC-SIGN se une de forma específica a carbohidratos con alto contenido en manosa (Jiménez D 2005), que contienen dos motivos  $\text{Man}\alpha 1-3$  [ $\text{Man}\alpha 1-6$ ] Man (Fig.1.8A1). Para demostrar cual de ellos participa en la unión a la lectina se hicieron experimentos por superposición de las regiones equivalentes de ambos oligosacáridos. Se observó que si interaccionara con la lectina el primer motivo trimanosa (recuadro negro Fig.1.8A1) el residuo N-acetilglucosamina unido en  $\beta 1-4$  chocaría con el residuo  $\text{Phe}^{325}$ . Sin embargo, si la interacción se lleva a cabo con el segundo motivo trimanosa (recuadro rojo en Fig.1.8A1), no hay efectos estéricos que impidan la unión. Por lo tanto, aunque el CDR de DC-SIGN reconoce trisacáridos unidos en conformación  $\text{Man}\alpha 1-3$  [ $\text{Man}\alpha 1-6$ ] Man, sólo puede llevarse a cabo la unión cuando la manosa central está unida en configuración  $\alpha$ -anomérica, estructura que se encuentra exclusivamente en oligosacáridos con alto contenido en manosa (Feinberg H 2001).



## **CAPITULO 2: OBJETIVOS**



## **2. OBJETIVOS**

1. Estudio de la contribución de las N-glicosilaciones en el plegamiento y función de los receptores de la familia ICAM.
2. Contribución de las N-glicosilaciones en el reconocimiento de los ICAMs por la lectina DC-SIGN.
3. Expresión y cristalización de fragmentos solubles del receptor ICAM-5.
4. Determinación de la estructura tridimensional de variantes de ICAM-5 con los dominios implicados en interacciones homofílicas.
5. Caracterización de las interacciones de adhesión homofílicas mediadas por el receptor ICAM-5.





## **CAPITULO 3: MATERIALES Y MÉTODOS**

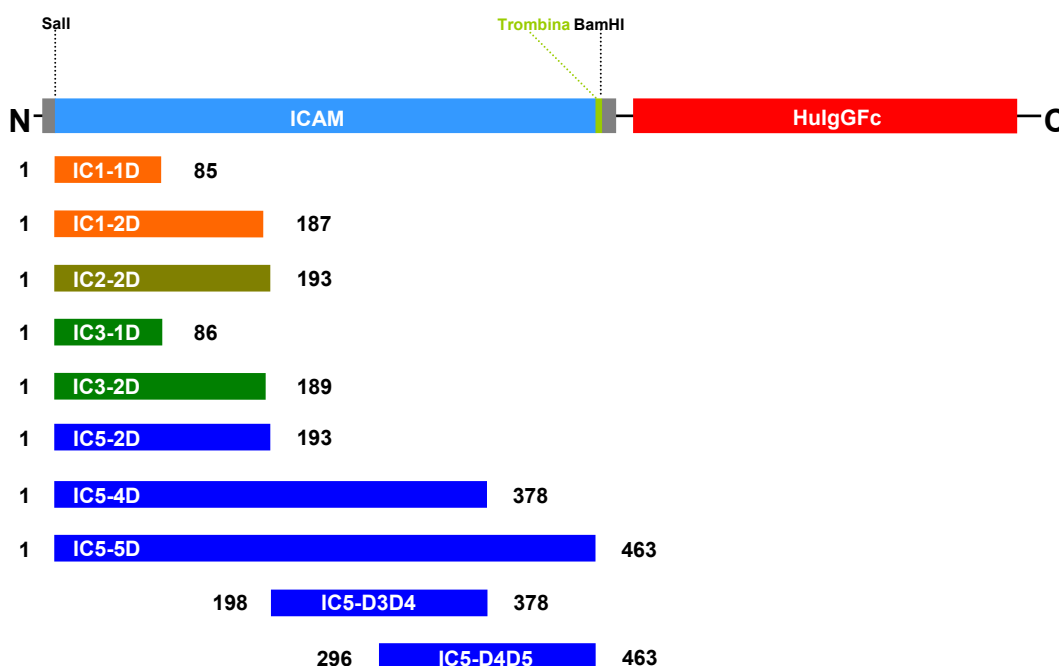


### 3. MATERIALES Y MÉTODOS

#### 3.1. Plásmidos y mutagénesis

Para llevar a cabo el presente trabajo se prepararon diferentes construcciones de proteínas de fusión clonando el ADN codificante para los dominios Ig de los ICAMs seguido de una señal de *splicing* y del ADN genómico de la región bisagra y Fc de la IgG1 humana (Fc) (Fig.3.1). En el extremo 3' del ADN codificante de los ICAMs se introdujo un sitio de corte de la proteasa trombina, para poder separar la proteína de la Fc en el proceso de purificación (Fig.3.1). Se prepararon las construcciones IC1-1D (residuos 1-85), IC1-2D (residuos 1-187), IC2-2D (residuos 1-193), IC3-1D (residuos 1-86) y IC3-2D (residuos 1-189) (Fig.3.1). Los ADNc mutantes de los ICAMs se prepararon mediante la técnica de solapamiento por PCR usando PfuI polimerasa (Stratagene) y los ADNc clonados en el sitio único XbaI del vector pAprM9 (Casasnovas JM 1999).

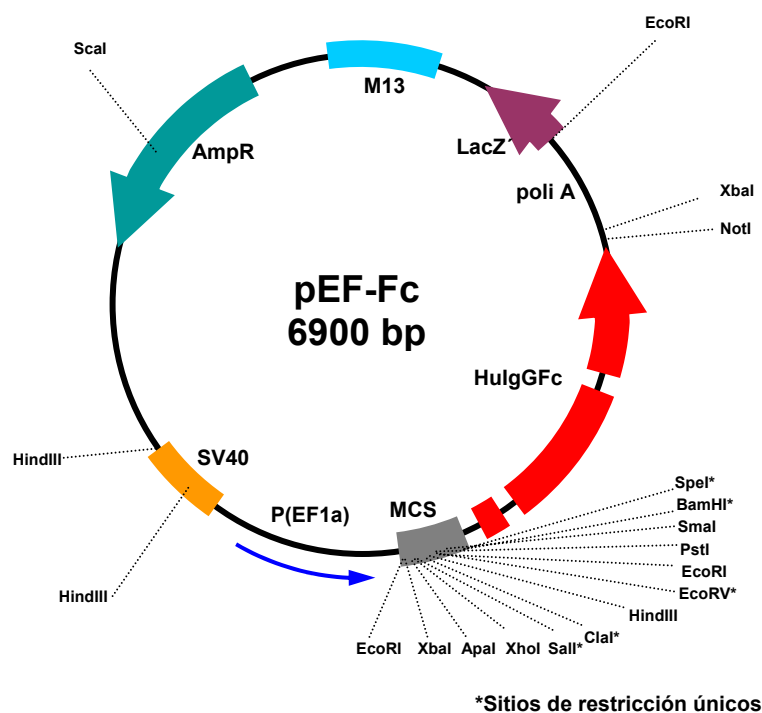
Para la realización del trabajo estructural con ICAM-5 se prepararon construcciones con diferentes dominios de la proteína, IC5-2D (residuos 1-193), IC5-4D (residuos 1-378), IC5-5D (residuos 1-463), IC5-D3D4 (residuos 198-378) e IC5-D4D5 (residuos



**Fig.3.1. Diferentes construcciones de ICAM-Fc** llevadas cabo en el presente trabajo, se indican los residuos que tiene cada una. Se introdujo un sitio de corte de la proteasa trombina en el extremo 3' del ADNc de cada ICAM. Los mutantes de ICAM-1 e ICAM-2 se clonaron de la misma manera.

296-463) (Fig.3.1).

Tanto fragmentos con diferentes dominios de ICAMs mutados o nativos se subclonaron en los sitios Sall y BamHI del vector pEF-Fc (Fig.3.2), vector usado para la expresión transitoria de proteínas en células de mamífero, construido a partir del vector pEF-BOS (Mizushima S 1990).



**Fig.3.2. Plásmido pEF-Fc.** Esquema del vector pEF-Fc, construido a partir del pEF-BOS (Mizushima S 1990). Contiene la zona de origen de SV40, el promotor humano EF-1 $\alpha$ , un sitio de poliadenilación, las regiones LacZ' y M13 y el gen de resistencia a la ampicilina. Como característica de este vector destaca la presencia del ADN de la región bisagra y Fc de IgG1 humana (HuIgGFC, rojo), también tiene una región de corte para múltiples sitios de restricción (MCS, gris), en cuyos sitios Sall-BamHI se clonaron las diferentes construcciones de ICAMs utilizados en este trabajo.

Para expresar la lectina DC-SIGN se utilizó el ADN codificante para el dominio lectina de DC-SIGN (Fig.1.7 B), que está formado por los residuos comprendidos entre Glu<sup>250</sup> y Leu<sup>392</sup>. Este ADN se clonó en el vector pET-27b (+) (NOVAGEN) en los sitios de restricción NdeI y XhoI.

### 3.2. Preparación de proteínas de fusión con Fc solubles

Las proteínas (ICAM-Fc) se expresaron de forma transitoria en células 293T transfectadas mediante el método del fosfato cálcico y con DMEM +5% suero (DuBridge RB 1987; Pear WS 1993). Los sobrenadantes celulares que contenían las proteínas se recogieron tres días después de la transfección. La secreción de proteínas al sobrenadante celular se monitorizó y cuantificó usando un ELISA (*enzyme-linked immunosorbent assay*) sándwich con anticuerpos contra la región Fc de las proteínas (DAKO). Se utilizó una proteína con Fc purificada y de concentración conocida como estándar para calcular la concentración de las proteínas en los sobrenadantes. La concentración de los ICAMs obtenida fue, aproximadamente, de 20 a 100 µg/ml.

### 3.3. Preparación de DC-SIGN

El dominio lectina de DC-SIGN se expresó utilizando un sistema bacteriano. La proteína se expresó formando cuerpos de inclusión en células de *Escherichia coli* BL21. Los agregados se solubilizaron en urea bajo condiciones desnaturalizantes (25 mM MES, pH 6.5, 8 M urea, 10 mM EDTA, 1 mM ditiotritol) y se sometieron a replegamiento *in vitro* mediante dilución 1:100 en un tampón que contenía 0.1 M Tris-HCL, pH 8.5, 400 mM L-arginina, 2 mM EDTA, 6.3 mM cisteamina, 3.7 mM cistamina y 0.1 mM fenilmetilsulfonil fluoruro. La solución de proteína (1 l) se concentró hasta 3 ml y se pasó por una columna de exclusión molecular Superdex-75 (Amersham Biosciences) en tampón 10 mM Tris, pH 8.0, con 100 mM NaCl (Fig.4.5A). La proteína DC-SIGN replegada que eluyó con el volumen de exclusión adecuado a su tamaño se purificó con una columna de intercambio iónico (Hi-Trap Mono-Q, Amersham Biosciences), eluyendo con una concentración salina de 250 mM NaCl en tampón Tris, pH 8.0 (Fig.4.5B). El correcto plegamiento de la proteína se corroboró por ensayos de unión al anticuerpo monoclonal anti DC-SIGN MR-1 (Relloso M 2002).

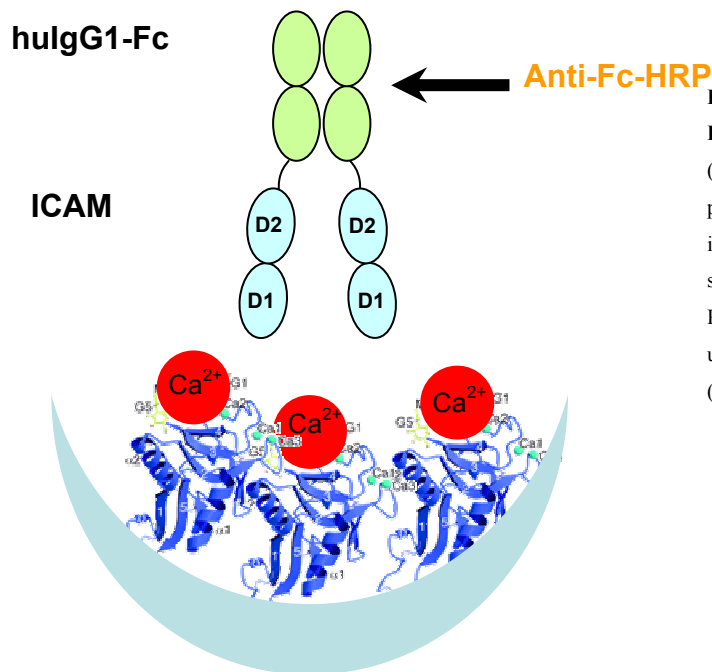
### 3.4. Unión de ICAMs a DC-SIGN

Para analizar la unión de los ICAMs a la lectina DC-SIGN diseñamos un ensayo con DC-SIGN unido al plástico y sobrenadantes celulares que contenían los ICAM-Fc (Fig.3.3). Para ello se usaron placas de 96 pocillos (Maxi Sorp, Nunc). Se cargaron los pocillos con 50 µl de DC-SIGN (10 µg/ml) en 20 mM Tris-HCl, 100 mM NaCl, pH 7.5, con 2.5 mM CaCl<sub>2</sub> (tampón de unión) y se incubó la placa toda la noche a 4° C. Al día siguiente se incubó la placa 1h a 37° C, se lavó 4 veces con tampón de unión, y se bloqueó con 2% BSA durante 1h a 37° C. Los sobrenadantes que contenían los ICAM-Fc se diluyeron en 1% BSA en tampón de unión en un rango de concentración entre 25 y 1 µg/ml, se añadieron a la placa y se dejaron incubarse 1h a 37° C. Tras lavar, se añadió anticuerpo anti-Fc marcado con HRP (*horseradish peroxidase*) (DAKO), diluido 1:500 en tampón de unión con 1% BSA. La presencia de HRP se detectó mediante el uso de OPD (DAKO), que se cuantificó mediante densidad óptica a 490 nm usando un lector de placas *Benchmark reader* (Bio-Rad). Como controles negativos se incluyeron pocillos sin DC-SIGN o sin ICAM-Fc, a los que se añadieron el resto de los reactivos.

### 3.5. Tratamiento de ICAMs-Fc con endoglicosidasa H y análisis de las proteínas

Se añadió endoglicosidasa H (*Endo-H*, New England Biolabs) o tampón de la enzima (control), a 200 µl de sobrenadante con ICAMs-Fc a una concentración de 20 µg/ml. La mezcla se incubó a 30° C toda la noche. Los ensayos de unión a DC-SIGN se hicieron como se describe arriba (Fig.3.3).

Los cambios de tamaño en las proteínas debidos a la hidrólisis de azúcares se analizaron por electroforesis y *Western blot* (Figs.4.9 y 4.11). Las muestras tratadas con *Endo-H* se inmunoprecipitaron a 4° C con proteína-A-sefárosa (Amersham Biosciences), se corrieron en un gel de acrilamida del 10% (SDS-PAGE) en condiciones reductoras y se transfirieron a una membrana *Hybond-P* (Amersham Biosciences). La membrana se bloqueó toda la noche en 5% de leche desnatada en polvo y 0.1% (v/v) de *Tween-20* en PBS. Posteriormente la membrana se incubó con anticuerpo anti-Fc marcado con HRP, y la unión del anticuerpo se detectó por quimioluminiscencia usando el sistema de detección ECL (Amersham Biosciences). El



**Fig.3.3. Ensayo de unión DC-SIGN-ICAMs.** Se cargaron placas de 96 pocillos (Maxisorp, Nunc) con DC-SIGN purificada, en presencia de calcio. Tras incubarla toda la noche a 4°C se añadieron sobrenadantes que contenían ICAMs-Fc. Para revelar se usó anticuerpo anti-Fc unido a HRP (*horseradish peroxidase*) (DAKO).

porcentaje de proteína sensible a *Endo-H* se determinó usando un lector *Fluor-S Multimager* (Bio-Rad).

### 3.6. Expresión de ICAMs en la superficie celular y ensayos de unión con anticuerpos, LFA-1 y HRV

Se expresaron proteínas silvestres y mutantes de ICAM-1 e ICAM-2 en la superficie celular. Para ello se transfectaron células COS, entre el 40 y el 60% de confluencia, en placas Petri de 100 mm usando el método de dextrano-DEAE (Casasnovas JM 1999). Cada experimento se llevó a cabo usando entre 0.1 y 8.0 µg de ADN silvestre y 6 µg de ADN mutante. Tres días después de la transfección las células se despegaron de las placas con 5 mM EDTA en PBS, se lavaron dos veces con medio L15, 2.5% FCS (tampón de unión) y se resuspendieron en el mismo a una concentración de  $2 \times 10^6$  células /ml.

La unión de anticuerpos a las moléculas de ICAM unidas a la superficie celular se detectó por inmunofluorescencia indirecta y citometría de flujo. Alrededor de  $1 \times 10^5$  células se incubaron con anticuerpos anti-ICAM a una concentración de 10 µg/ml. El porcentaje de células que expresaban los epítomos de los anticuerpos se determinó tras

sustraer la señal de fondo proveniente de células control transfectadas con pAprM9 (Casasnovas JM 1999).

Se llevaron a cabo experimentos de unión con LFA-1 purificada y pegada a plástico (Casasnovas JM 1999). Se usó LFA-1 inmunopurificada a 0.5 mg/ml diluida 100 veces en tampón TSM (20 mM Tris, 150 mM NaCl, 2 mM MgCl<sub>2</sub>, pH 9.0). Se añadieron 30 µl por pocillo en una placa de 96 de fondo plano y se incubó a 4°C toda la noche. La placa se bloqueó añadiendo a cada pocillo 200 µl de 2% BSA en tampón de unión, durante 30 minutos a 37°C y se lavó dos veces con tampón de unión antes de llevar a cabo el ensayo de unión. Las células COS transfectadas se marcaron con el marcador fluorescente BCECF (Molecular Probes). Se tomaron las células a una concentración aproximada de  $7 \times 10^4$  en 25 µl de tampón de unión y añadieron por triplicado a los pocillos que contenían LFA-1 pegada y 75 µl de tampón de unión. La placa se centrifugó a 1000 rpm durante 1 minuto y se incubó a 37°C durante 30 minutos. Se midió la fluorescencia antes y después de quitar las células no unidas por medio de tres o cuatro lavados con 100 µl de tampón de unión por pocillo a temperatura ambiente, para posteriormente calcular los porcentajes de células transfectadas con moléculas de ICAM silvestres y mutantes unidas a LFA-1. La unión específica se calculó restando la señal de fondo provocada por las células control transfectadas con el vector vacío pAprM9.

Los ensayos de unión de rinovirus a células transfectadas se hizo con HRV3 y HRV14 marcado con <sup>35</sup>S (Martin S 1993). Las células transfectadas ( $1 \times 10^5$  células por pocillo) con ADN recombinante silvestre (2-8 µg) y mutante (4 µg) fueron sembradas 48h después de la transfección en placas de 96 pocillos y cultivadas 24h a 37°C. Se diluyó [<sup>35</sup>S]HRV3 o HRV4 ( $2 \times 10^4$  cpm) en 80 µl de PBS, 5% FCS, se añadió por duplicado a los pocillos y se incubó a 35°C durante 1h. En los experimentos de bloqueo con anticuerpos, las células se preincubaron durante 15 minutos con el anticuerpo anti-IC1 RR1/1 a una concentración de 30 µg/ml (Staunton DE 1990). La solución con el virus se aspiraba y los pocillos se lavaban dos veces con 50 µl de PBS, 5% FCS. Las células en contacto con los virus se recogieron con 30 µl de Tritón X-100 al 1% y se hizo un recuento de centelleo. Las cuentas obtenidas de los virus unidos a las células se normalizaron sustrayendo la unión de fondo provocada por células control. La tasa de unión se determinó con las cuentas obtenidas de células que expresaban en superficie cantidades similares de ICAM-1 silvestre y mutante, determinado con el anticuerpo anti ICAM-1 CL203 (Staunton DE 1990).



### **3.7. Resolución de la estructura tridimensional de IC5-4D**

#### **3.7.1. Preparación y purificación de IC5-4D**

Se prepararon varias construcciones con diferentes dominios funcionales: D1-D2 (IC5-2D), D1-D4 (IC5-4D), D1-D5 (IC5-5D), y también IC5-D3D4 y IC5-D4D5 (Fig.4.12) como se indica en la figura 3.1.

Inicialmente las proteínas se expresaron de forma transitoria en células 293T y posteriormente se prepararon líneas celulares que secretaban la proteína de interés al sobrenadante celular de forma estable. Estos clones celulares se obtuvieron a partir de células CHO-Lec 3.2.8.1 (Stanley 1989), transfectadas por el método del fosfato cálcico y usando el sistema de expresión de la glutamina sintasa (Casasnovas JM 1998) .

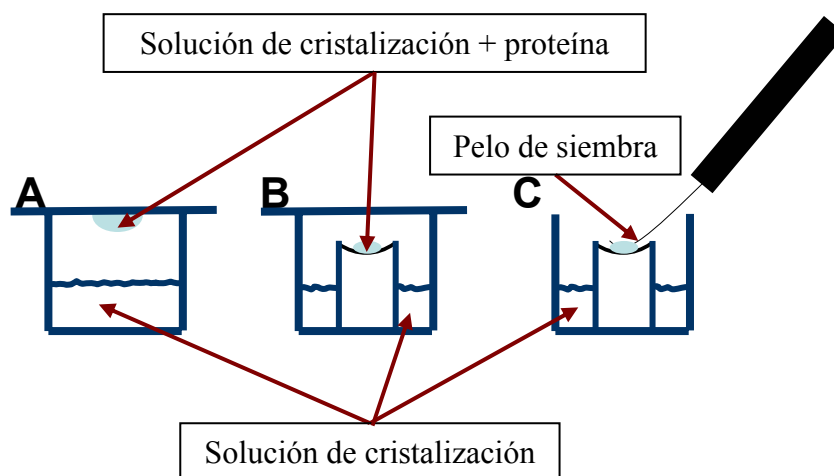
El primer paso de la purificación de las proteínas de fusión consistió en una cromatografía de afinidad a través de una columna de proteína-A-sefarosa (Amersham Biosciences), que presenta una alta afinidad por la región Fc. La proteína se eluyó con bajo pH (50 mM glicina, 150 mM NaCl, pH 3.0) e inmediatamente se neutralizó con 1M tris pH 8.0. Las fracciones que contenían proteína se concentraron hasta un volumen de 3.5 ml y se digirieron con trombina (Amersham Biosciences) 16 horas a 30° C (Fig.4.13A). Mediante un nuevo paso por una columna de proteína-A-sefarosa (Amersham Biosciences), se separó el fragmento de IC5 de la Fc (Fig.4.13B). Seguidamente la proteína de interés se purificó por exclusión molecular mediante una columna de Superdex-200 (Amersham Biosciences) en tampón 10 mM HEPES, 100 mM NaCl, pH 7.5 (Fig.4.13C, D).

Para la proteína IC5-4D, incluida en este trabajo, las fracciones provenientes de la columna de exclusión se sometieron a un último paso de purificación a través de una columna de intercambio iónico SP (-) (Amersham Biosciences). La proteína se eluyó mediante un gradiente salino desde 100 mM a 1M NaCl en tampón 10 mM HEPES, pH 7.5 (Fig.4.13E, F).

#### **3.7.2. Concentración y cristalización de IC5-4D**

La proteína purificada se concentró por centrifugación usando un concentrador Microsep 10K (Pall), hasta una concentración aproximada de 20 mg/ml.

Para la cristalización de la proteína se utilizaron los métodos de difusión de vapor por *gota colgante* y *gota sentada* (Fig.3.4). La técnica de difusión de vapor consiste en mezclar una pequeña muestra de proteína con un volumen igual, o similar de solución de cristalización (tampón, sal y precipitante). La gota se sella en un reservorio que contiene entre 500 y 1000  $\mu\text{l}$  de solución de cristalización. La diferencia de concentración entre la gota y el reservorio conduce al sistema hacia el equilibrio por difusión a través de la fase de vapor. Si las condiciones son buenas, la proteína llega a un estado supersaturado que hace que los cristales empiecen a formarse cuando la gota y el pocillo alcanzan el equilibrio (Bergforss 1999).



**Fig.3.4. Técnicas de cristalización.** **A. Gota colgante**, la mezcla de proteína y solución precipitante se hacía sobre un cubreobjetos siliconizado al que posteriormente se daba la vuelta y se sellaba en un reservorio de una placa de cultivos de 24 pocillos. **B. Gota sentada**, se usaron placas *cryschem*, posteriormente se sellaban los pocillos con cinta adhesiva. **C. Siembra**, se introducían pequeños cristales, “semillas”, que actuaban como puntos de nucleación. Para ello se utilizaba un pelo de gato que se pasaba por la solución que contenía las “semillas” y posteriormente por una gota equilibrada, dejando en ésta varios puntos de nucleación a partir de los cuales nacían nuevos cristales.

En el método de *gota colgante* (Fig.3.4A) se mezcló 1.5  $\mu\text{l}$  de solución de cristalización con 1.5  $\mu\text{l}$  de proteína (a una concentración de 15 mg/ml) sobre un cubreobjetos siliconizado con el que se selló el pocillo que contenía la solución precipitante en una placa de cultivo de 24 pocillos. Para el método de *gota sentada* (Fig.3.4B) se utilizó una placa *Cryschem*. Las placas se dejaron reposar a una temperatura de 22° C, hasta que el conjunto alcanzó el equilibrio por difusión de vapor (Bergforss 1999).

Se obtuvieron tres formas cristalinas (Fig.4.14), siendo necesario utilizar la técnica de siembra para la reproducción de uno de los cristales (Bergforss 1999) (Fig.3.4C). Este método consiste en introducir en una gota equilibrada pequeños cristales (semillas)

que actúan como puntos de nucleación a partir de los cuales crecerán nuestros cristales. Para llevar a cabo el método se usaron gotas en placas de *gota sentada* un día después de su preparación.

### 3.7.3. Colección y procesamiento de datos cristalográficos

Cada conjunto de datos se colectó a partir de un monocristal congelado bajo nitrógeno líquido a  $-160^{\circ}\text{C}$ . Los cristales recogidos en la solución de cristalización se dializaron contra la solución de congelación, que contenía etilenglicol al 25% como crioprotector (Rodgers 1994).

Los datos de difracción se colectaron usando radiación de sincrotrón y de difractómetros convencionales. Se procesaron usando los programas MOSFLM (Leslie 1992) y XDS (Kabsch 1993). El objetivo de los programas de procesamiento de datos es generar, a partir de un conjunto de imágenes de difracción, los valores de los índices ( $hkl$ ) con sus respectivas intensidades, junto con una estimación inicial de los valores de la celda unidad. Posteriormente los datos se escalaron usando el programa SCALA (Collaborative Computational Project 1994). Las intensidades de difracción medidas en un experimento de difracción no están en la misma escala, ya que son afectadas por diferentes factores físicos del experimento, la mayoría de los cuales son difíciles de medir directamente. El proceso de escalado pone todas las difracciones medidas en una escala común ajustándolas a un modelo. Esto produce un conjunto de datos internamente consistente, analizando la coincidencia de reflexiones equivalentes tras escalar se puede determinar la calidad de los datos (Evans 2006).

### 3.7.4. Resolución y refinamiento de la estructura

Cuando un cristal se somete a difracción de rayos X, el resultado obtenido es un patrón de difracción, el cual consiste en millones de señales que constituyen el dato bruto. Cada una de las señales está determinada por tres parámetros que son necesarios para obtener el mapa de densidad electrónica. Dichos parámetros son la posición, la intensidad y la fase de las ondas que forman cada señal. Los dos primeros son relativamente fáciles de determinar, sin embargo la resolución de la fase constituye otro

de los grandes inconvenientes de la cristalografía de rayos X (Anexo I). Con el fin de solventar este problema, frecuentemente se emplea la técnica de reemplazamiento isomorfo. Este método se basa en la incorporación de metales pesados en la proteína cristalizada sin modificar su estructura. Pero no siempre se puede utilizar este método, por lo que habitualmente se utiliza otro llamado reemplazo molecular. Este método aprovecha la similitud estructural compartida por los miembros de una familia de proteínas, de forma que si se conoce la estructura de uno de los miembros de la familia se puede determinar la posición y la orientación de la molécula desconocida dentro de la celda unidad, y de esta manera resolver su estructura (Anexo I).

El resultado de una buena cristalización y determinación de la fase es una imagen tridimensional formada por nubes de electrones de la proteína que se conoce con el nombre de mapa de densidad electrónica. A partir de este mapa y con la secuencia de aminoácidos, se construye un modelo inicial de la estructura de la proteína, el cual da lugar a un modelo final tras ser sometido a una serie de refinamientos. Este último modelo de la estructura de la proteína consiste en las coordenadas cartesianas de cada uno de los átomos de la molécula excepto de los átomos de hidrógeno. Una propiedad importante del modelo final es su resolución, es decir, la mínima distancia que puede ser definida. La unidad de este parámetro es el Amstrong (Å), y cuanto menor sea su valor mayor será la resolución, es decir, el mapa de densidad electrónica será de mayor calidad.

En un primer momento se trabajó en la resolución de la estructura de IC5-4D mediante el método de reemplazamiento isomorfo (MIR). Para ello se colectaron diferentes cristales “nativos” y “derivados”, incubados en distintas soluciones de átomos pesados, variando sus concentraciones y tiempos de incubación. Las posiciones de los átomos pesados en la celda se determinaron por diferencias de Patterson y de Fourier, y los sitios fueron refinados utilizando el programa MLPHARE (Collaborative Computational Project 1994). Se calcularon mapas de densidad electrónica usando el programa FFT (Collaborative Computational Project 1994), mapas que se sometieron a modificación de densidad con el programa DM (Collaborative Computational Project 1994).

Para la determinación de la estructura de IC5-4D por el método de reemplazo molecular se usaron los programas AMORE y PHASER (Collaborative Computational Project 1994), así como la estructura del dominio 1 de ICAM-2 (Casasnovas JM 1997) y las estructuras de los dominios 2, 3 y 4 de ICAM-1 (Casasnovas JM 1998), (Yang Y

2004). Con el programa PHASER se obtuvo una solución en la que los cuatro dominios aparecían concatenados, tal y como se espera. Estos dominios se sometieron a “*rigid body refinement*” con el programa REFMAC (Collaborative Computational Project 1994) y se calcularon unas fases iniciales a partir de reemplazo molecular. Estas fases permitieron identificar algunos sitios de unión de átomos pesados determinados por el método de reemplazo isomorfo, lo que corroboró los resultados obtenidos por ambos métodos. Las fases obtenidas con los átomos pesados calculadas con el programa MLPHARE se combinaron con las de reemplazo molecular usando el programa SIGMAA (Collaborative Computational Project 1994). Se generaron nuevos mapas de densidad electrónica a 3 Å de resolución, mejorados por técnicas de “modificación de densidad” usando DM (Collaborative Computational Project 1994). Con este mapa se construyó manualmente un modelo inicial de IC5-4D, usando el programa O (Jones TA 1991). El modelo se ha sometido a refinamiento con el programas CNS (Brünger AT 1998), extendiéndose la resolución hasta 2.5 Å.



## **CAPITULO 4: RESULTADOS**

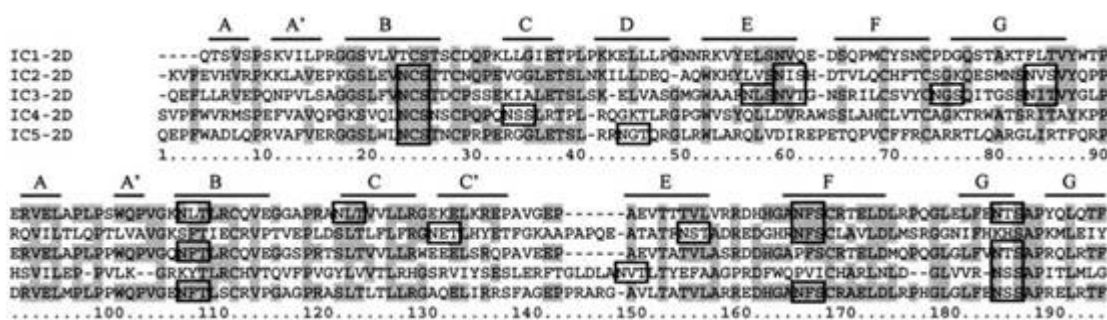




## 4. RESULTADOS

### 4.1. Contribución de los glicanos en el plegamiento y función de los ICAMs

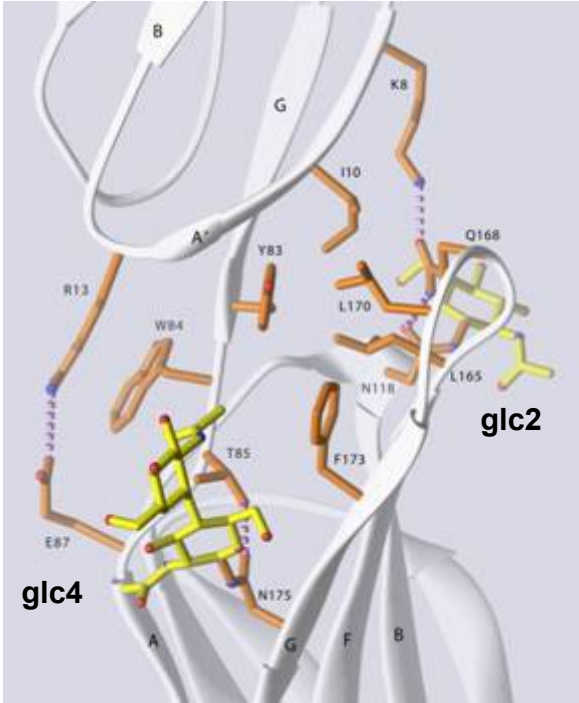
La resolución de las estructuras cristalinas de ICAM-1 (Casasnovas JM 1998) e ICAM-2 (Casasnovas JM 1997) aportaron nuevos datos para entender la importancia de las glicosilaciones en la función de los ICAMs. Los dos dominios N-terminal de los ICAMs humanos están altamente glicosilados, con un número de azúcares que oscila entre los cuatro de ICAM-1 e ICAM-4 a los siete de ICAM-3 (Fig.4.1). El primer y el último azúcar, situados en la hebra B del dominio 1 y en la hebra G del dominio 2, respectivamente, están altamente conservados.



**Fig.4.1. Alineamiento y sitios de N-glicosilación en los dominios N-terminales de ICAMs.** Alineamiento de los dos dominios N-terminal (2D) de los ICAMs humanos (de IC1 a IC5) con los sitios de glicosilación recuadrados. Los alineamientos fueron hechos con ClustalW y Modeler. El primer dominio (D1) está en la parte de arriba y el segundo dominio (D2) en la de abajo. Las líneas indican los residuos incluidos en las láminas-β indicadas en cualquiera de las dos estructuras de ICAM-1 e ICAM-2.

El azúcar situado en la hebra G del segundo dominio de ICAM-1 (glc4, Fig.1.6 B, Fig.4.2) corre paralelo al bucle FG en la estructura de ICAM-1 y participa de la región entre los dominios 1 y 2 (Fig.4.2). La primera N-acetilglucosamina se empaqueta junto a la cadena lateral de la Thr<sup>85</sup>, como se ha visto en todas las estructuras disponibles (Bella J 1998; Casasnovas JM 1998; Shimaoka M 2003). También se ha observado un contacto entre la segunda N-acetilglucosamina y los residuos Tyr<sup>83</sup> y Phe<sup>173</sup> (Casasnovas JM 1998; Shimaoka M 2003) (Fig.4.2). Este azúcar es el que tiene el factor de temperatura más bajo (Tabla1) y por lo tanto el que presenta menor flexibilidad, quizá por su interacción con los residuos proteicos. La eliminación por mutagénesis de este azúcar (glc4) redujo la expresión de una molécula soluble con los dos dominios N-

terminal de ICAM-1 y también modificó la unión a LFA-1 (Staunton DE 1990; Bella J 1998).



**Fig.4.2. Región entre los dominios 1 y 2 de ICAM-1.** Se muestran detalle de las interacciones entre los residuos que forman parte de la unión entre los dominios 1 y 2. La primera N-acetilglucosamina del azúcar situado en la hebra G del segundo dominio (glc4), se empaqueta junto a la cadena lateral de la Thr85. La segunda N-acetilglucosamina contacta con los residuos Tyr83 y Phe173. Estas interacciones pueden ser la causa del bajo factor de temperatura de esta glicosilación comparado con el resto de glucanos unidos a ICAM-1 (Tabla1).

ICAM-1		ICAM-2		
Azúcar	B-Factor	Azúcar	B-Factor	% Accesible
glc1	85	glc1	54	5.96
glc2	81	glc2	69	8.89
glc3	65	glc3	61	6.28
glc4	50.5	glc4	69	12.35
		glc5	76	11.78
		glc6	64	9.66

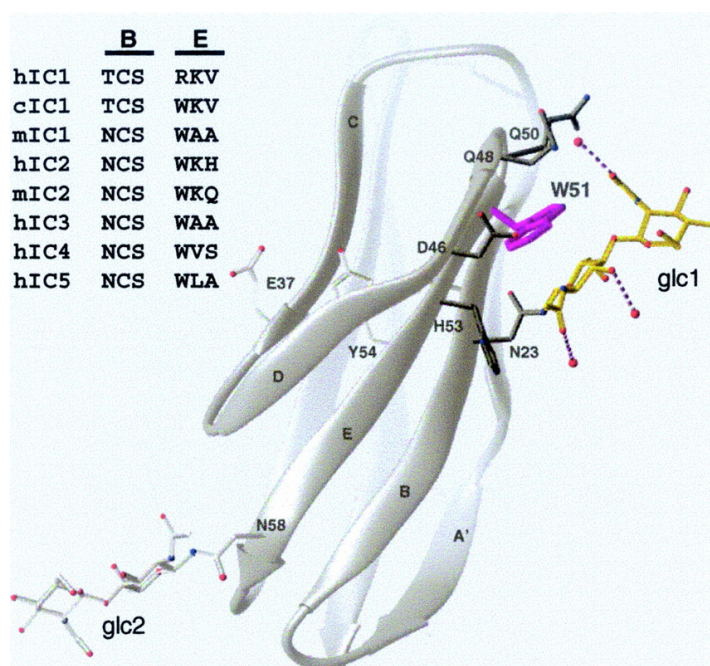
**Tabla 1.** Factores de temperatura (B) de la primera N-acetilglucosamina de cada azúcar de ICAM-1 e ICAM-2. También se indica el porcentaje expuesto al solvente de los azúcares unidos a ICAM-2. (Los azúcares se numeran igual que en la figura 1.6).

4.1.1. Un motivo proteína-glicano conservado en el primer dominio de los ICAMs

La estructura de ICAM-2 reveló algunas peculiaridades para el azúcar (glc1) altamente conservado en el primer dominio de los ICAMs (Fig.1.7, Fig.4.1). Se localiza en el centro del dominio inmunoglobulina, al contrario que el resto de azúcares que están en la parte inferior de los dominios 1 y 2 (Fig.1.7 A). De todos los

azúcares de ICAM-2 es el que tiene el factor de temperatura más bajo (Tabla 1), lo que indica una baja flexibilidad provocada por la interacción con un residuo aromático muy conservado (Fig.4.3).

La estructura muestra un contacto entre los primeros residuos de glc1 y el residuo conservado Trp<sup>51</sup> (Fig.4.3), con la primera N-acetilglucosamina descansando sobre el anillo aromático de la cadena lateral del triptófano. De forma que la accesibilidad al solvente de los átomos que forman el primer residuo del carbohidrato es significativamente menor que la de otros residuos unidos al resto de sitios de glicosilación (Tabla 1). Asimismo el grupo metilo de la primera N-acetilglucosamina descansa a aproximadamente 4 Å del anillo del residuo His<sup>53</sup> (Fig.4.3). Debido a estos contactos, glc1, en lugar de proyectarse perpendicular al eje del dominio, corre paralelo al mismo hacia la parte superior del mismo. Trp<sup>51</sup> se encuentra rodeado por el azúcar por un lado y por la cadena lateral alifática de Gln<sup>48</sup> y Gln<sup>50</sup> por el otro, preservando al anillo aromático del contacto con el solvente.



**Fig.4.3. Motivo azúcar-triptófano conservado en el dominio de unión a LFA-1 en ICAM-2.** Modelo del dominio N-terminal de ICAM-2 con el azúcar glc1 (amarillo) y el Trp<sup>51</sup> (magenta). Los residuos que rodean al Trp<sup>51</sup> están coloreados en gris oscuro, y el residuo crítico en la unión a LFA-1, Glu<sup>37</sup> en blanco. Los átomos de oxígeno y nitrógeno están coloreados en rojo y azul, respectivamente. Las moléculas de agua unidas por puentes de hidrógeno a glc1 se representan como esferas rojas. La tabla incluida representa el motivo azúcar-triptófano en los ICAMs humanos (hIC1 a hIC5), ICAM-1 de chimpancé (cIC1) e ICAMs de ratón (mIC1,2).

El motivo azúcar-triptófano está altamente conservado en la subfamilia de los ICAMs (Fig.4.3 izquierda). Todos estos receptores excepto ICAM-1 de humano y chimpancé conservan el motivo, por lo que se deduce que debe tener una función conservada en la subfamilia. La mutación tanto del azúcar como del triptófano en

ICAM-2 reduce significativamente tanto la expresión en superficie del receptor como la unión a LFA-1 (ver punto 4.1.3 abajo). Se observó un fenotipo similar para ICAM-3 humano (Klickstein LB 1996), indicando que el motivo azúcar-triptófano es fundamental para la correcta conformación del dominio de unión a la integrina en la subfamilia de los ICAMs. Asimismo, una molécula mutante de ICAM-1 que contenía el motivo azúcar-triptófano (T20N/R49W) se expresó en la superficie celular de forma parecida al ICAM-1 silvestre, como se pudo comprobar mediante unión al anticuerpo monoclonal CL203, que reconoce el dominio 4 (Tabla 2). La unión a LFA-1 fue prácticamente igual en el ICAM-1 silvestre que en el mutante, lo que indicaba el correcto plegamiento del dominio para el reconocimiento de la integrina. La introducción por mutagénesis en ICAM-1, tanto del azúcar como del triptófano no tuvo un efecto significativo en la expresión, aunque el mutante T20N se unió deficientemente a la integrina (Tabla 2).

Mutante	LFA-1	HRV	CL203	RR1/1	R6.5	LB2	YH370	7F7
AprM9	0	0	0	0	0	0	0	0
ICAM-1	100	100	100	100	100	100	100	100
+IC1-Ab	22 (15)	3 (4)						
T20N	59 (12)	78 (11)	101 (23)	87 (5)	90 (9)	62 (6)	58 (8)	0 (0)
T20A		123 (5)	82	115	110	100		
R49W	94 (38)	92 (7)	96 (38)	10 (3)	100 (3)	98 (4)	96 (16)	102 (11)
T20N-R49W	110 (9)	46 (11)	102 (34)	103 (6)	104 (7)	24 (3)	39 (8)	5 (7)

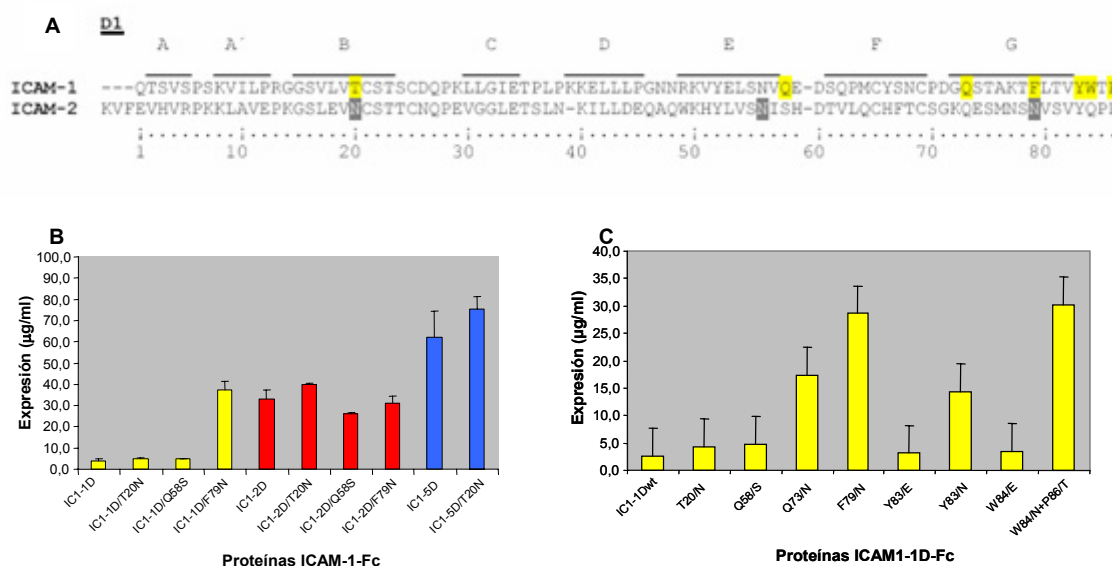
**Tabla 2.** Actividad de unión de mutantes de glicosilación de ICAM-1 a LFA-1, HRV y anticuerpos. Unión de LFA-1, HRV y anticuerpos a mutantes de ICAM-1 expresados en la superficie celular. La unión se normalizó con la obtenida de células transfectadas con el receptor silvestre, determinada con el anticuerpo monoclonal CL203 (materiales y métodos). Los valores de unión de los mutantes y el silvestre al anticuerpo CL203 se determinaron a partir del porcentaje de células que expresaban el epítipo, transfectadas con la misma cantidad de ADN mutante o silvestre. Se incluye la unión de LFA-1 y HRV en presencia del anticuerpo anti-ICAM-1 RR1/1 (+IC1-Ab). Los resultados se obtuvieron de la media de al menos tres experimentos, (excepto para el mutante T20A), con la desviación estándar indicada en paréntesis.

El sitio de unión al anticuerpo R6.5, que incluye epítopos de los dominios 1 y 2 de ICAM-1 (Staunton DE 1990), no se ve afectado por la adición del motivo azúcar-triptófano en el dominio N-terminal (Tabla 2). La inclusión del azúcar (mutante T20N) bloqueaba la unión del anticuerpo 7F7 y reducía de forma parcial la unión de los anticuerpos LB2 e YH370. Sin embargo, la introducción de la doble mutación (T20N-R49W) disminuyó de forma cooperativa la unión de estos dos anticuerpos (Tabla 2), a pesar de la ausencia de fenotipo para la mutación (R49W). Estos resultados sugieren un cambio en la conformación del azúcar. Así, en ausencia del residuo triptófano el azúcar se debe disponer perpendicular al dominio, de forma

similar a glc2 de ICAM-2 (Fig.4.3). Sin embargo, la orientación del azúcar en el doble mutante debe ser similar a la descrita por la estructura tridimensional de ICAM-2 para glc1 (Fig.4.3), cubriendo los epítomos LB2 e YH370. Adicionalmente se podría esperar que el mutante de ICAM-1 con el motivo azúcar-triptófano exhibiese el azúcar decorando la zona de unión a rinovirus humano (HRV), situada en la parte superior del primer dominio inmunoglobulina (Kolatkár PR 1999), lo que podría afectar la unión del virus al receptor. Se observó que la inclusión del azúcar en la zona media del primer dominio inmunoglobulina de ICAM-1 (mutante T20N) disminuía levemente la unión del virus, mientras que la adición del triptófano (mutante T20N-R49W), reducía significativamente la interacción (Tabla 2).

#### 4.1.2. Contribución de los glicanos al plegamiento y expresión del dominio N-terminal de ICAM-1 aislado

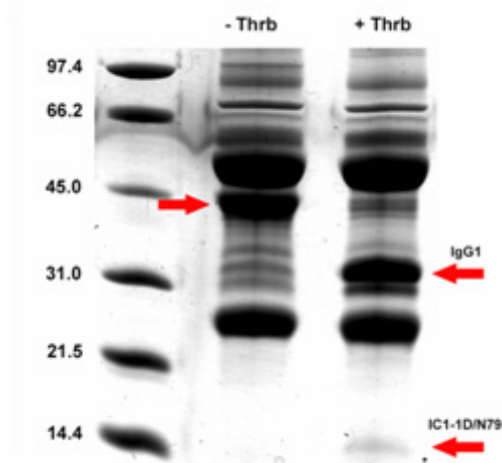
Para estudiar la contribución de los carbohidratos al plegamiento y la expresión de



**Fig.4.4. Mutantes de glicosilación de IC1-1D-Fc y expresión de los mismos.** A. Alineamiento de los residuos correspondientes al primer dominio de ICAM-1 e ICAM-2. Los residuos de ICAM-1 mutados se indican en amarillo, mientras que los sitios de N-glicosilación de ICAM-2 se muestran en gris. B. Proteínas silvestres y mutantes se expresaron en células 293-T, la concentración de las proteínas se monitorizó mediante ELISA-sandwich usando anticuerpos anti-Fc. Se observa expresión de diferentes mutantes y pérdida de expresión del silvestre con un solo dominio N-terminal, salvo para el mutante IC1-1D/F79N, que se expresa a un nivel similar a las proteínas con dos dominios. C. Comparación de los niveles de expresión entre diferentes mutantes de ICAM-1 con un solo dominio, se observa como al introducir azúcares en la parte final del dominio, (hebras F, G), los mutantes se expresan de forma similar a moléculas de ICAM-1 con sus dos primeros dominios inmunoglobulina.

los dominios Ig en la subfamilia de los ICAMs, analizamos el efecto de la introducción de N-glicosilaciones en el primer dominio de ICAM-1. Se seleccionó este modelo porque hasta la fecha no se había podido expresar el primer dominio de ICAM-1 aislado y porque este dominio carece de glicosilaciones (Fig.1.7, Fig.4.1). Si bien es posible expresar el dominio N-terminal de ICAM-3 aislado. Por ello, nos pareció interesante intentar expresar un solo dominio de ICAM-1 introduciendo N-glicosilaciones en diferentes partes de la molécula.

Para ello, se prepararon diferentes mutantes mediante la técnica de solapamiento por PCR (materiales y métodos). Se introdujeron residuos de Asn en las posiciones homólogas en que se encuentran los sitios de glicosilación en ICAM-2 (Fig.4.4A). El ADNc que codifica para el primer dominio del receptor se clonó en el vector pEF-Fc con objeto de producir moléculas solubles en las que el primer dominio estaría fusionado a la región Fc de IgG1 humana, con un sitio de reconocimiento de trombina entre ambas proteínas (materiales y métodos). Con los vectores recombinantes se transfectaron células 293-T y se analizó y cuantificó la secreción de las proteínas a los sobrenadantes celulares mediante ensayos ELISA con anticuerpos anti-Fc, (materiales y métodos). Los mutantes de glicosilación que tienen los dos primeros dominios N-



**Fig.4.5. Comprobación del primer dominio del mutante IC1-1D/F79N.** Electroforesis de la proteína de fusión parcialmente purificada tratada (+Thr) o sin tratar (-Thr) con trombina, cuyo sitio de reconociendo se encuentra entre el dominio 1 de ICAM-1 y la región Fc de IgG1. Gel del 12% acrilamida bajo condiciones reductoras. El tamaño del marcado molecular (kDa) se incluye a la izquierda.

terminales de ICAM-1 se expresaron de forma similar a las moléculas silvestres (Fig.4.4B). Sin embargo no se observó expresión de las moléculas de fusión que contenían solamente el primer dominio silvestre (Fig.4.4B, C). Dos mutantes con glicosilaciones localizadas en las hebras  $\beta$  B (N20) y E (N58) presentaron niveles de expresión muy bajos, mientras que la introducción de glicosilaciones en la hebra G o en la región C-terminal de dominio incrementó la expresión de las proteínas con un único dominio. Los mutantes IC1-1D/F79N y W84N+P86T se expresaron a un nivel similar a las proteínas con dos dominios (Fig.4.4B, C).

Seguidamente analizamos la presencia del primer dominio en la proteína de fusión. Se expresó el mutante IC1-1D/N79 en célula CHO-K1 por medio del sistema de la glutamina sintetasa (materiales y métodos). Se obtuvieron cinco clones celulares que expresaban la proteína de fusión, que se secretaba al medio extracelular. La expresión se monitorizó por ensayo de ELISA. La proteína se purificó usando una columna de proteína-A-sefarosa y posteriormente se analizaron por electroforesis muestras de proteína tratadas o sin tratar con trombina (Fig.4.5). Se observó que la proteína de fusión de aproximadamente 45 kDa (1D + Fc) desaparecía por el tratamiento con la proteasa, generando proteínas con el tamaño esperado para la región Fc y el dominio N-terminal glicosilado de ICAM-1 (IC1-1D/F79N) (Fig.4.5).

#### **4.1.3. Importancia de las N-glicosilaciones en la expresión de ICAM-2 y en el reconocimiento de LFA-1**

La contribución de las N-glicosilaciones en la expresión de ICAMs y la unión a LFA-1 ha sido estudiada por mutagénesis con ICAM-1 e ICAM-3 (Staunton DE 1990; Klickstein LB 1996), pero no con ICAM-2. Los azúcares unidos al primer dominio de ICAM-2 están conservados en ICAM-3, mientras que los del segundo dominio no lo están (Fig.4.1). Para estudiar la contribución de los azúcares en la expresión de la proteína en la membrana celular y en el reconocimiento de la integrina, se prepararon mutantes de ICAM-2 que carecían diferentes sitios de N-glicosilación. Las proteínas se expresaron en la superficie celular (materiales y métodos) y la expresión se monitorizó usando cuatro anticuerpos monoclonales (Tabla 3). La supresión en ICAM-2 del azúcar conservado (glc1) unido al residuo Asn<sup>23</sup> con tres mutaciones diferentes (N23A, N23T y S25A) se tradujo en un descenso muy acusado tanto de la expresión como de la unión a LFA-1 (Tabla 3). El efecto fue incluso más acusado que el que se observó al mutar el azúcar homólogo en ICAM-3 (Klickstein LB 1996). La eliminación de cualquiera de los otros azúcares en ICAM-2 tuvo escaso efecto en la unión a LFA-1, que estuvo relacionada con los niveles de expresión (Tabla3). La doble mutación de azúcares en el segundo dominio generó un descenso significativo en la expresión del receptor (Tabla 3), mientras que la expresión en superficie de un triple mutante fue prácticamente indetectable.



Mutante	LFA-1	IC2/1	IC2/2	BT-1	6D5
Aprm9	0	0	0	0	0
ICAM-2	100	100	100	100	100
+IC2-Ab	5.4 (7)				
+LFA-Ab	2.5 (5)				
$\Delta$ glc1 (N23A)	1.5 (2)	25 (14)	2 (2)		37 (22)
$\Delta$ glc1 (N23T)	3.4 (3)	26 (8)	6 (6)	3 (1)	34 (4)
$\Delta$ glc1 (S25A)	1.1 (1)	23 (16)	3 (0)	5 (4)	44 (2)
W51A	1.5 (2)	31 (2)	4 (4)	2 (0)	3 (6)
$\Delta$ glc2 (S60A)	139.3 (32)	65 (22)	56 (30)	64 (20)	74 (34)
$\Delta$ glc3 (S83A)	74.1 (14)	90 (16)	82 (16)	84 (18)	104 (44)
$\Delta$ glc4 (T131A)	125.6 (2)	84 (24)	84 (14)	67 (11)	47 (0)
$\Delta$ glc5 (T154A)	76.5 (18)	85 (13)	73 (27)	78 (16)	59 (29)
$\Delta$ glc6 (S165A)	66.5 (6)	69 (13)	65 (10)	57 (15)	67 (25)
$\Delta$ glc4,5	44.9 (24)	51 (10)	54 (12)	46 (10)	52 (2)
$\Delta$ glc5,6	10.0 (2)	12 (4)	20 (5)	8 (3)	38 (2)

**Tabla 3.** Unión de LFA-1 y anticuerpos a mutantes de ICAM-2 expresados en la superficie celular. La unión de LFA-1 se normalizó con un silvestre expresado en la misma cantidad, (materiales y métodos) (Casasnovas JM 1999). La expresión de las proteínas en la superficie celular se determinó con el anticuerpo CBR-IC2/1. La unión de los anticuerpos a los mutantes se normalizó con el porcentaje de células que expresaban ICAM-2 silvestre transfectadas con la misma cantidad de ADN. Se incluyen los datos de unión en presencia de los anticuerpos anti-ICAM-2, CBR-IC2/2 (+IC2-Ab), y anti-LFA-1, TS1/18 (+LFA-Ab). Se muestra la media de tres experimentos con las desviaciones estándar en paréntesis. Los azúcares eliminados ( $\Delta$ ) se numeraron igual que en la figura 1.6, y los sitios mutados se muestran en paréntesis.

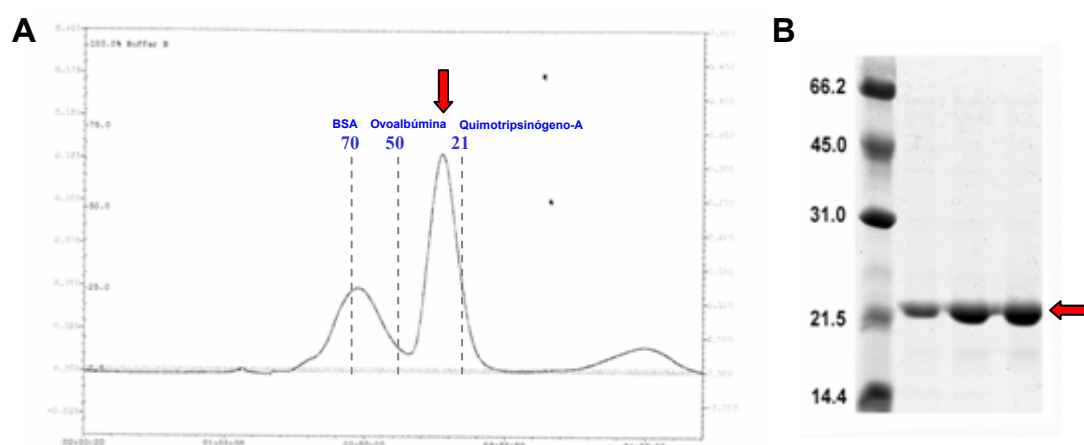
## 4.2. Contribución de las N-glicosilaciones al reconocimiento de ICAMs por la lectina DC-SIGN

### 4.2.1. Preparación del dominio de lectina de DC-SIGN

DC-SIGN (*dendritic-cell-specific ICAM-3 grabbing non-integrin*) es una lectina de tipo C presente en la superficie de las células dendríticas que juega un papel importante en el proceso de captura y presentación de antígeno (Geijtenbeek TB 2000). Es una proteína de membrana de tipo II, cuyo dominio extracelular está formado por 23 residuos en tandem repetidos siete veces de forma completa y una de forma incompleta (Figdor CG 2002). Posee un dominio de reconocimiento de carbohidratos dependiente de  $\text{Ca}^{2+}$  en el extremo COOH terminal (Fig.1.8B) (Feinberg H 2001). Este dominio adopta la conformación típica de las lectinas de tipo C (Drickamer 1999).

Expresamos el dominio de lectina (residuos Glu<sup>250</sup>-Leu<sup>392</sup>) de DC-SIGN usando la cepa BL21 de *Escherichia coli*. Inicialmente la proteína se expresó como cuerpos de inclusión, que se sometieron a replegamiento *in vitro* (materiales y métodos). Tras el replegamiento, la proteína se purificó usando una columna de exclusión Superdex-75 (Amersham Biosciences). La lectina eluyó con un volumen de exclusión correspondiente a 20 KDa, que coincide con el peso molecular de la proteína



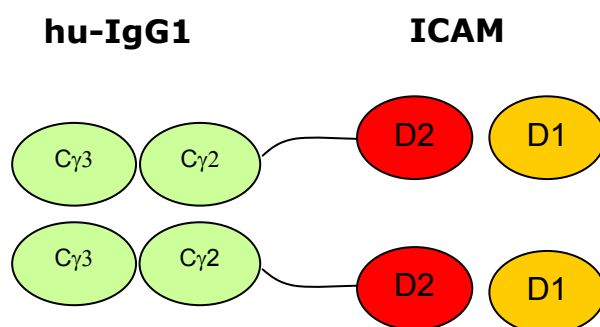


**Fig.4.6. Purificación de DC-SIGN.** **A**, Cromatograma indicando la absorbancia a 280 nm de las fracciones procedentes de la columna de exclusión molecular Superdex-75 (Amersham Biosciences). Las líneas discontinuas en azul indican el volumen de exclusión de proteínas estándar con su tamaño en KDa. **B**, Gel electroforesis donde se analizó la pureza de la proteína eluida de la columna de intercambio iónico. Las flechas indican posición de la proteína en la cromatografía y electroforesis.

(Fig.4.6A). Seguidamente, la lectina se purificó por cromatografía de intercambio iónico usando una columna Hi-Trap Mono-Q (materiales y métodos) y se obtuvo una proteína de elevada pureza (Fig.4.6B). Para comprobar que la proteína estaba correctamente plegada y era funcional se hicieron ensayos de unión al anticuerpo monoclonal anti DC-SIGN MR-1 (Relloso M 2002), observando reconocimiento de la proteína por el anticuerpo.

#### 4.2.2. Glicosilación en ICAMs y su contribución al reconocimiento por DC-SIGN

Los dos dominios N-terminal de todos los ICAMs humanos están altamente glicosilados. El número de sitios de glicosilación varía desde los 4 de ICAM-1 a los 7

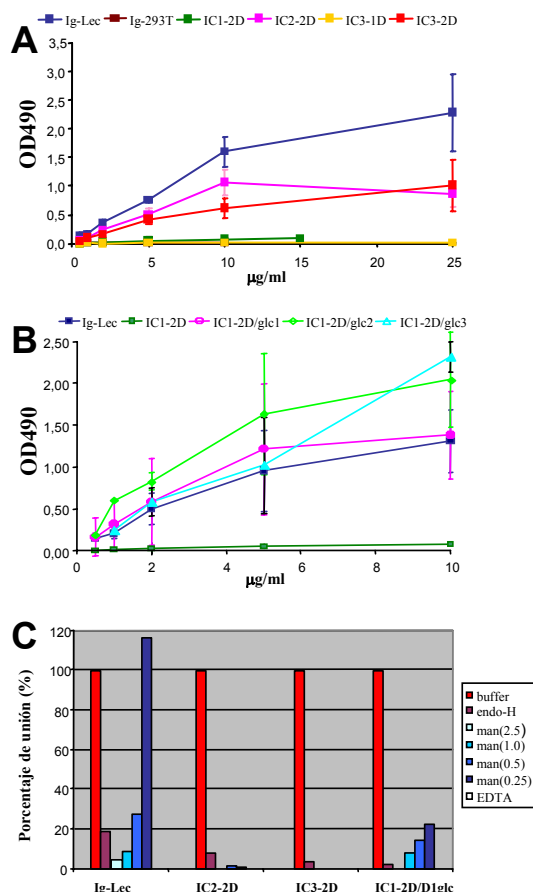


**Fig.4.7. Construcción de las proteínas de fusión ICAMs-Fc.** Las proteínas se prepararon clonando el ADN codificante para los dominios Ig de los ICAMs seguido del ADN genómico de la región bisagra y Fc de IgG1 humana (materiales y métodos).

de ICAM-3 (Fig. 4.1).

La lectina DC-SIGN se une específicamente y con gran afinidad a los dos dominios N-terminales de ICAM-2 e ICAM-3 en presencia de calcio, sin embargo no se ha detectado unión a ICAM-1 (Figdor CG 2002). Para definir los determinantes de esta especificidad diseñamos un ensayo (Fig.3.3) con DC-SIGN inmovilizada sobre plástico y moléculas de ICAM solubles con sus dos dominios N-terminal fusionados a la región Fc de IgG1 (Fig.4.7). La integridad funcional de los ICAMs en las proteínas de fusión se comprobó usando anticuerpos monoclonales anti-ICAM (no mostrado). Como control positivo se usó una proteína consistente en los dos dominios de inmunoglobulina N-terminales de Nectin-1 y que contenía carbohidratos ricos en manosa (Ig-Lec). Como control negativo usamos la misma proteína con glicanos complejos y expresada en células 293-T, que no se une a DC-SIGN a pesar de tener 5 N-glicosilaciones.

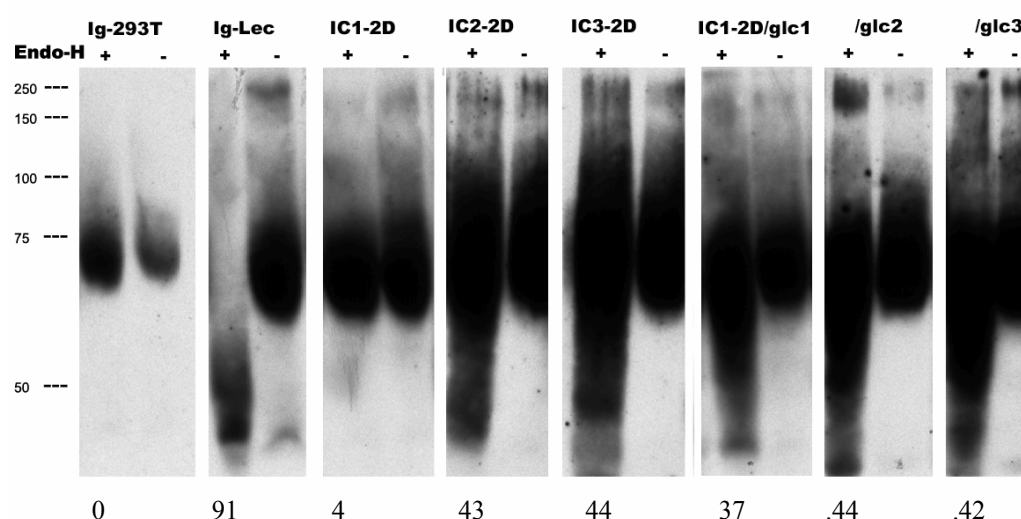
Moléculas solubles expresadas en células 293-T y con los primeros dos dominios N-terminales de ICAM-2 (IC2-2D) e ICAM-3 (IC3-2D) se unieron a DC-SIGN (Fig.4.8A), mientras que moléculas con el dominio homólogo de ICAM-1 (IC1-2D) o



**Fig.4.8. Unión de ICAMs a DC-SIGN.** A. Proteínas expresadas en células 293-T que contienen los dominios 1 y 2 de ICAM-1 (IC1-2D), ICAM-2 (IC2-2D), ICAM-3 (IC3-2D) o el dominio 1 de ICAM-3 (IC3-1D). Los ensayos de unión se hicieron en paralelo con la proteína Nectin-1 expresada en células CHO-Lec (Ig-Lec) o en 293-T (Ig-293T). B. Ensayos de unión con Ig-Lec, IC1-2D y mutantes de ICAM-1 introducidos en la posición 20 (IC1-2D/glc1), 56 (IC1-2D/glc2) ó 79 (IC1-2D/glc3). C. Especificidad en la unión de ICAMs a DC-SIGN. Los porcentajes de unión se calcularon a partir de la OD a 490 nm de proteínas tratadas y no tratadas. En los ensayos de unión a DC-SIGN se usaron las proteínas indicadas a 10 µg/ml tratadas con tampón o con Endo-H, o en presencia de concentraciones crecientes de mannan o EDTA (25mM). No se detectó unión en las muestras tratadas con EDTA. Se muestra la media y la desviación estándar de tres experimentos independientes.

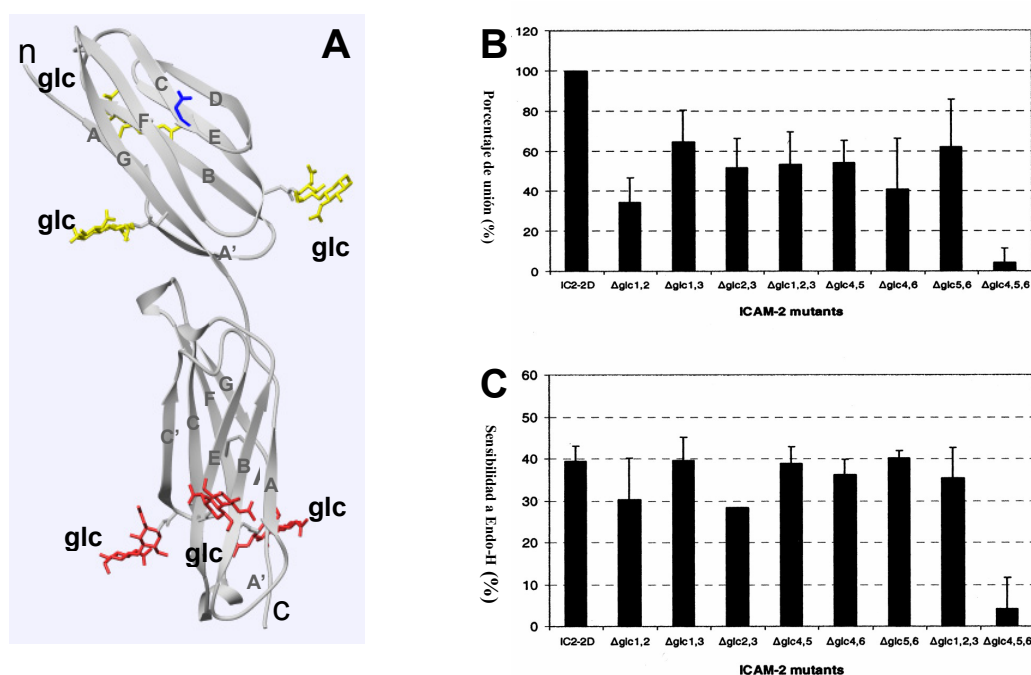
un solo dominio N-terminal de ICAM-3 (IC3-1D) no se unieron a la lectina. Sin embargo, moléculas de ICAM-1 con dos dominios y con una N-glicosilación adicional en tres sitios diferentes del primer dominio se unieron a DC-SIGN (Fig.4.8B). La unión de DC-SIGN a los ICAMs se inhibía mediante quelación del calcio por EDTA, con tratamiento con Endo-H o con alta concentración de mannan (Fig.4.8C), lo cual sugería que DC-SIGN reconoce azúcares con alto contenido en manosa o de tipo híbrido.

Para entender mejor la unión de los ICAMs a DC-SIGN se analizaron por electroforesis las proteínas tratadas con endo-H (Fig.4.9). En las muestras de proteína Ig-Lec se observó una esperada reducción de tamaño debida a la exclusiva presencia de N-glicosilaciones con un alto contenido en manosas. Sin embargo no se observó reducción de tamaño en la misma proteína expresada en células 293-T (capaces de generar glicosilaciones complejas), indicando la ausencia de azúcares con alto contenido en manosas reconocidos por DC-SIGN. Un resultado similar se obtuvo con IC1-2D (Fig.4.9) y IC3-1D (no mostrado), proteínas que no unen a la lectina. Por el contrario, alrededor del 40% del resto de las proteínas mostraron reducción de tamaño tras el tratamiento con Endo-H, indicando la presencia de carbohidratos ricos en manosa responsables de la unión de estos receptores a la lectina (Fig.4.8A y C).



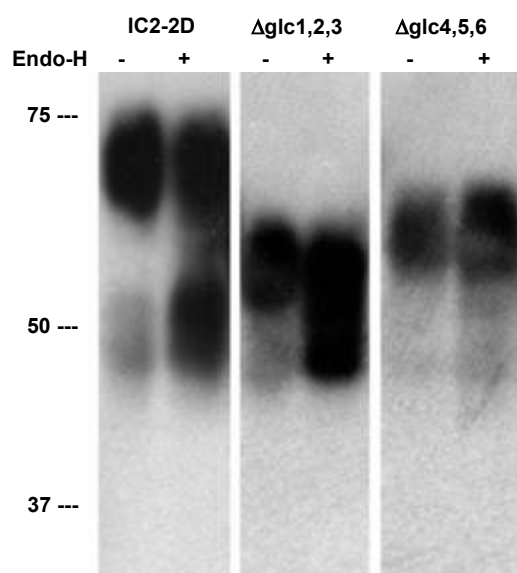
**Fig.4.9. Sensibilidad de ICAMs solubles a Endo-H.** *Western blotting* de las proteínas de fusión con Fc tratadas (+) o no tratadas (-) con Endo-H. Para ello se incubaron los sobrenadantes que contenían cada una de las proteínas a una concentración de 20 µg/ml por 16 hr. a 30°. Posteriormente las proteínas se inmunoprecipitaron y se sometieron a *Western blot*, revelando con un anticuerpo anti-Fc marcado con HRP. Abajo se indica el porcentaje cada proteína sensible a Endo-H.

Seguidamente determinamos los sitios de glicosilación de ICAM-2 que contienen carbohidratos implicados en la unión a DC-SIGN. Este receptor contiene 6 carbohidratos, 3 en cada uno de los dominios, 5 de ellos localizados en la parte inferior de los dominios Ig (Fig.4.10A). Se llevaron a cabo mutaciones selectivas de los diferentes sitios y posteriormente se analizó la capacidad de unión a la lectina de los mutantes. La eliminación de un solo sitio de glicosilación no afectó a la unión a DC-SIGN (no mostrado), mientras que la mutación de dos sitios presentó una disminución en la afinidad, aunque el porcentaje de unión continuaba siendo significativo (Fig.4.10 B). La eliminación de todos los sitios de glicosilación del primer dominio mostró una reducción de la unión en un 50%, mientras que la supresión de los tres sitios del segundo dominio produjo una pérdida total de unión (Fig.4.10B). Estos resultados indicaban que los azúcares con alto contenido en manosa sensibles a Endo-H (Fig.4.10C) sólo están ausentes en los mutantes de ICAM-2 que carecen de N-glicosilaciones en el segundo dominio, mostrando que deben ser los responsables de la actividad de unión a la lectina.



**Fig.4.10. Reconocimiento de los mutantes de glicosilación de ICAM-2 por DC-SIGN y sensibilidad a Endo-H.** A. Modelo con los dos dominios Ig de ICAM-2 coloreados en gris, las N-glicosilaciones del primer y segundo dominio están pintadas en amarillo y rojo, respectivamente. B. Actividad de unión a la lectina de los mutantes de ICAM-2 sin los azúcares indicados, eliminados por mutagénesis. Los porcentajes de unión se calcularon a partir de las OD490 obtenidas de los ensayos de unión con DC-SIGN sobre plástico y proteínas silvestres y mutantes solubles. El resultado es la media de cinco experimentos independientes. C. Porcentaje de cada muestra de proteína sensible al tratamiento con endo-H, determinado por *Western blotting*. El resultado es la media de dos experimentos.

El menor tamaño del mutante  $\Delta\text{glc1,2,3}$  comparado con  $\Delta\text{glc4,5,6}$ , (Fig.4.11) sugiere un menor tamaño de los azúcares unidos al segundo dominio, lo cual indica glicosilaciones más simples como las que contienen un alto contenido en manosa. Como prueba adicional se sometieron las proteínas a digestión con Endo-H y se observó que el mutante  $\Delta\text{glc1,2,3}$  era sensible a la proteasa (Fig.4.11), lo que demuestra que posee azúcares con alto contenido en manosa, mientras que el tamaño del mutante  $\Delta\text{glc4,5,6}$  no se afectó por la digestión con Endo-H (Fig.4.11). Esto explica también la ausencia de unión del mutante  $\Delta\text{glc4,5,6}$  a la lectina DC-SIGN. La pérdida de unión a DC-SIGN de IC3-1D (Fig.4.8A) sugiere también que los azúcares con alto contenido en manosa deben estar en el segundo dominio de ICAM-3.

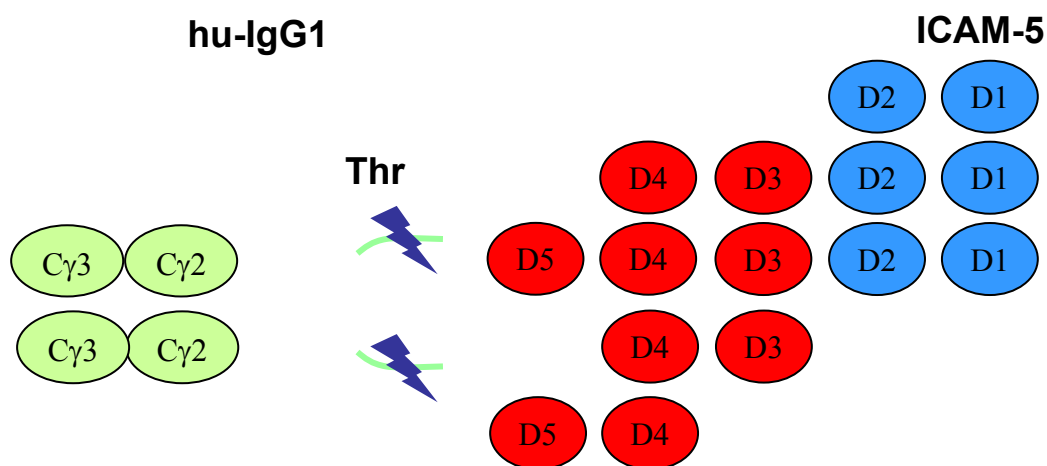


**Fig.4.11. Diferencia de tamaño entre los mutantes de ICAM2.** *Western blotting* donde se muestra la diferencia de tamaño entre moléculas de IC2-2D silvestre y mutantes que carecen de glicosilaciones en el primer ( $\Delta\text{glc1,2,3}$ ) o el segundo dominio ( $\Delta\text{glc4,5,6}$ ), así como su sensibilidad a Endo-H. El hecho de que el mutante sin glicosilaciones en el segundo dominio sea más grande indica que los azúcares eliminados son más simples y por tanto de menor tamaño, como los de alto contenido en manosa. Los cuales son además sensibles a Endo-H, como se puede ver en el mutante que tiene azúcares en el segundo dominio ( $\Delta\text{glc1,2,3}$ ), cosa que no ocurre con el mutante ( $\Delta\text{glc4,5,6}$ ). Esto explica también la ausencia de unión del mutante  $\Delta\text{glc4,5,6}$  a la lectina DC-SIGN.

### 4.3. Estructura de ICAM-5

#### 4.3.1. Preparación de fragmentos de ICAM-5 con los dominios extracelulares implicados en adhesión homofílica

Los 5 dominios N-terminales de ICAM-5 parecen estar implicados en adhesión homofílica entre moléculas localizadas en células opuestas (ver introducción). Para la realización de estudios estructurales sobre ICAM-5 se prepararon varias construcciones que codificaban para diferentes dominios funcionales: D1-D2 (IC5-2D), D1-D4 (IC5-4D), D1-D5 (IC5-5D), y también IC5-D3D4 y IC5-D4D5, (Fig.3.1, Fig.4.12).

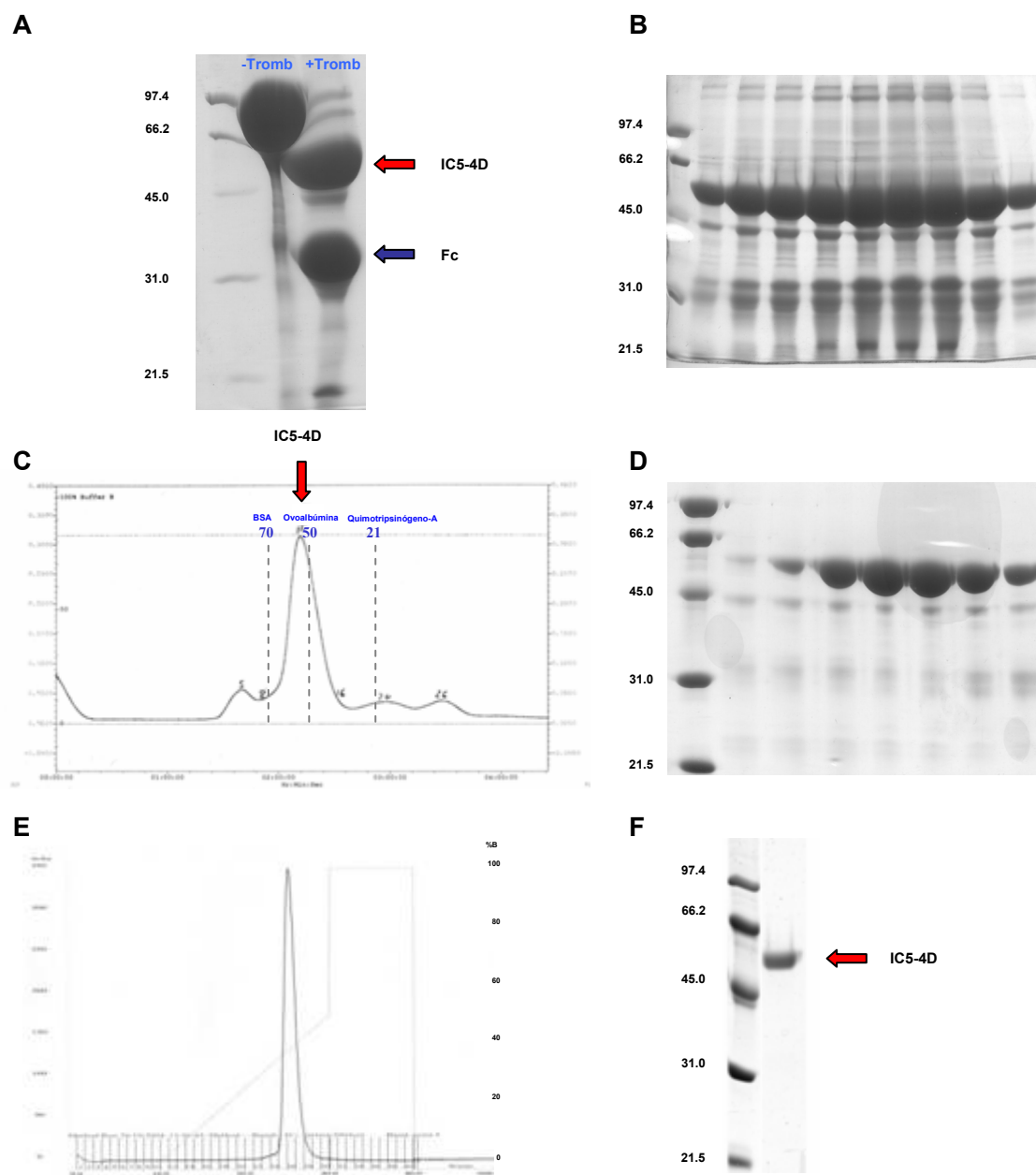


**Fig.4.12. Fragmentos solubles de ICAM-5.** Esquema de proteínas de fusión que contienen los dominios funcionales de ICAM-5 unidos a la región Fc de IgG1 humana. Para poder separar ambas proteínas durante el proceso de purificación, se introdujo un sitio de corte de la proteasa trombina.

Las proteínas se expresaron en células CHO-Lec usando el sistema de expresión de la glutamina sintasa (Casasnovas JM 1998). Se seleccionaron clones celulares que expresaban de forma estable cada una de estas proteínas, identificándose por medio de un ensayo de ELISA que reconoce la región Fc. La expresión de la proteína de fusión era de aproximadamente 4 mg por litro de sobrenadante. Inicialmente preparamos IC5-2D y IC5-4D.

Los intentos de cristalizar IC5-2D por otros miembros del grupo no dieron resultados satisfactorios, así que mi trabajo se centró en IC5-4D. Inicialmente IC5-4D se purificó a partir de sobrenadantes celulares por cromatografía de afinidad con una columna de proteína-A-sefarosa (Amersham Biosciences), que tiene gran afinidad por la Fc humana. Posteriormente la proteína de fusión se proteolizó con trombina para

separar IC5-4D de la Fc (Fig.4.13A). Seguidamente la muestra se sometió a una nueva cromatografía de afinidad con columna de proteína-A-sefarosa para retener la Fc (Fig.4.13B).



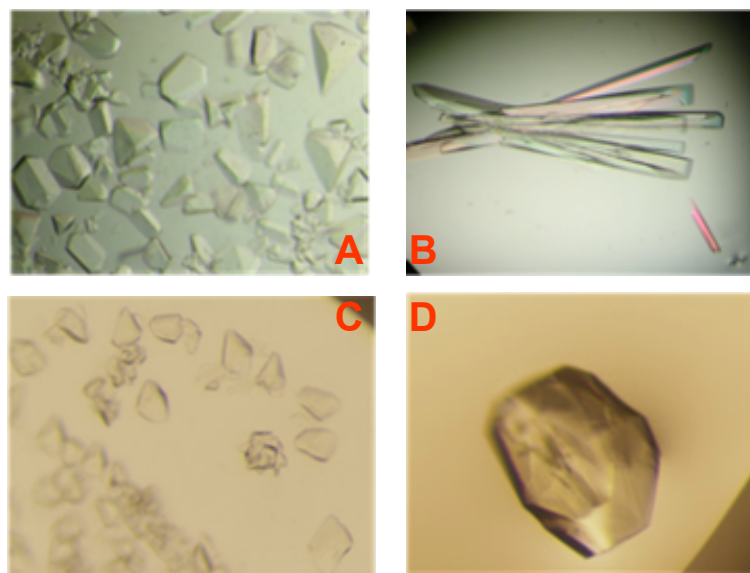
**Fig. 4.13. Purificación de ICAM5-4D.** A. Electroforesis donde se analizó la efectividad de la digestión con la proteasa trombina. En el primer carril se incluyó la proteína sin cortar y en el segundo a la proteína cortada. B. Electroforesis tras pasar la proteína cortada con trombina por una columna de proteína A con el fin de separarla de la Fc. C. Cromatograma indicando la absorbancia a 280 nm de las fracciones procedentes de la columna de exclusión de molecular Superdex-200 (Amersham Biosciences). Las líneas discontinuas en azul indican el volumen de exclusión de proteínas estándar, para las que se indican su tamaño en KDa. D. Electroforesis donde se analizaron las fracciones obtenidas en la columna de exclusión molecular. E. Cromatograma indicando la absorbancia a 280nm de las fracciones obtenidas en la columna de intercambio iónico SP(-). F. Análisis por electroforesis de la pureza de la proteína eluída de la columna de intercambio iónico.

A continuación la proteína se sometió a cromatografía de exclusión con una columna Superdex-200 (Amersham Bioscience), donde la muestra eluyó con un volumen de exclusión correspondiente a 50 KDa, que es el peso molecular esperado para esta proteína (Fig.4.13C y D). Esta muestra de proteína se sometió a un último paso de purificación a través de una columna de intercambio iónico SP (-) (Amersham Biosciences) (Fig.4.13D), tras el cual obtuvimos proteína con alta pureza (Fig.4.13F).

La proteína se concentró por centrifugación hasta una concentración aproximada de 20 mg/ml (materiales y métodos).

### 4.3.2. Cristalización de la proteína

A partir de muestras de IC5-4D purificada y concentrada llevamos a cabo ensayos de cristalización por el método de difusión de vapor con *gota colgante* y *gota sentada*, (materiales y métodos). Inicialmente se probaron diferentes cantidades de precipitante (polietilenglicol-4000, PEG-4K), diferentes condiciones de pH, tipo de tampón, concentración de proteína, adición de compuestos y diferentes temperaturas (4, 18 y 22° C).



**Fig.4.14. Cristales de IC5-4D obtenidos en diferentes condiciones.** A Cristales obtenidos con 10% PEG 4K , 100 mM NaAc pH 5.6, a 18°C (X1). B Cristales crecidos con 10% PEG 4K, 100 mM Tris pH 8.5, a 18°C (X2). C y D Cristales crecidos con 10% PEG 4K, 100 mM Tris pH 8.5, a 22°C por el método de siembra (X3).



Inicialmente con estos ensayos obtuvimos dos tipos de cristales diferentes (Fig.4.14) dependiendo del pH utilizado. Cristales tipo bloque en condiciones que contenían 10% PEG-4K, 100 mM NaAc pH 5.6, 18-22°C (Fig.4.14A, forma cristalina X1) y cristales tipo barra en las mismas condiciones pero con 100 mM Tris pH 8.5 (Fig.4.14B, forma cristalina X2). Ambos tipos de cristales crecían con una concentración de proteína entre 5 y 15 mg/ml, siendo esta última la concentración óptima. Estos cristales se dializaron frente a una solución idéntica a la de cristalización y que contenía etilenglicol 20-30 %. El etilenglicol actúa como crioprotectante que evita la rotura del cristal en el proceso de congelación (Rodgers 1994). Ninguno de estos dos cristales descritos difractaron a muy alta resolución, observando un patrón de difracción hasta aproximadamente 4 Å de resolución en la línea de rayos X ID14-2 del ESRF (european synchrotron radiation facility) (Tabla 4: 3,4).

Una nueva forma cristalina minoritaria se identificó en condiciones de alto pH. Estos cristales, a pesar de crecer a pH 8.5 tenían una forma tipo bloque e irregular (Fig.4.14C y D, forma cristalina X3) parecida a los obtenidos a pH 5.6. Cuando la proteína estaba recién preparada, en condiciones de alto pH sólo salían cristales de tipo barra, X2 (Fig.4.14B), sin embargo, si se dejaba la proteína más de una semana a 4° C y se repetían las condiciones de cristalización ya no salían los cristales del tipo X2 y en algunas gotas se podían ver estos nuevos cristales X3, pero la eficiencia del proceso de cristalización era muy baja y además no era fácil reproducir los resultados.

Para reproducir la preparación de los cristales se utilizó el método de *siembra* (materiales y métodos), mediante el cual conseguimos estandarizar la cristalización y obtener los cristales necesarios para llevar a cabo la resolución de la estructura, ya que estos cristales difractaban bastante bien logrando una resolución de 2.3 Å (Tabla 4: 1,2).

#### 4.3.3. Colección y procesamiento de datos cristalográficos

Los datos de difracción de los cristales se colectaron en las líneas del sincrotrón del ESRF. Cada conjunto de datos se colectó a partir de un monocristal congelado bajo nitrógeno líquido a -160° C. Los cristales fueron previamente dializados frente a una solución crioprotectora que contenía solución de cristalización con etilenglicol al 25%.

La mayor parte de los datos de difracción se colectaron usando radiación de sincrotrón (ESRF, Grenoble). Se procesaron usando los programas MOSFLM (Leslie

1992) y XDS (Kabsch 1993) y se escalan con SCALA (Collaborative Computational Project 1994) (materiales y métodos). En la tabla 4 se muestran las estadísticas obtenidas para las diferentes formas cristalinas.

Los cristales X1 pertenecen al grupo espacial R3 ( $a = 228.56$   $b = 228.56$   $c = 69.98$   $\alpha = \beta = 90.0$   $\gamma = 120.0$ ), tienen una molécula en la unidad asimétrica y alrededor de 65% de solvente. Los datos de difracción con los que se cuentan en la actualidad alcanzan una resolución de 4.2 Å. Los cristales X2 difractaron a una resolución similar, aunque pertenecen al grupo espacial P4<sub>3</sub>22 ( $a = 96.07$   $b = 96.07$   $c = 321.92$   $\alpha = \beta = \gamma = 90.0$ ), tienen dos moléculas en la unidad asimétrica y alrededor de 66% de solvente. Finalmente, la forma cristalina X3 difracta a alta resolución, pertenece al grupo espacial P21 ( $a = 76.59$   $b = 46.91$   $c = 95.79$   $\alpha = \gamma = 90.0$   $\beta = 104.34$ ) y tiene una molécula en la unidad asimétrica, con un contenido de solvente de aproximadamente 60%.

NATIVOS									
CRISTAL	LINEA	RES.	COMPL. (%)	REDUN.	I/ $\sigma$	R(%)	G E	MOL.U.A	%SOLVENTE
1 X1	ID14.2	25- 4.2	98.5	3.2	7.5 (2.5)	6.3 (30.1)	R3	1	65
2 X2	ID14.2	25- 4.4	99.2	4.3	6.0 (2.4)	7.7 (31.0)	P4 <sub>3</sub> 22	2	66
3 X3.1	ID14.2	20- 2.3	96.7	6.0	10.0 (3.3)	4.8 (23.2)	P21	1	60
4 X3.2	BM16	25- 2.5	99.6	3.7	7.7 (2.2)	6.3 (34.3)	P21	1	60

**Tabla 4.** Estadísticas de los datos de difracción colectados con cristales nativos pertenecientes a las tres formas cristalinas de IC5-4D. Los datos se procesaron con XDS (Kabsch 1993) y se escalan con SCALA (Collaborative Computational Project 1994), de donde proceden estas estadísticas. Se indica la línea del ESRF donde se colectaron, la resolución, completitud, redundancia, el valor de I sobre sigma, el factor R, grupo espacial, las moléculas por unidad asimétrica y la concentración de solvente.

#### 4.3.4. Determinación de la estructura de la forma X3

##### Método del reemplazamiento isomorfo múltiple (MIR)

Con el fin de obtener un conjunto de fases que nos permitiesen resolver la estructura de la proteína comenzamos utilizando el método del reemplazamiento isomorfo múltiple (MIR). El método consiste en introducir en la celda unidad átomos que sean grandes dispersores de los rayos X. La dificultad estriba en el hecho de que los átomos introducidos no deben distorsionar la estructura cristalina de la proteína nativa, es decir, que los cristales así obtenidos han de ser isomorfos a los nativos. Para llevar a cabo esta técnica hay que coleccionar datos a partir de cristales “nativos” y de cristales “derivativos”,

que se obtienen difundiendo complejos de metales pesados a través de los canales que poseen los cristales de proteína, y en donde normalmente existen cadenas laterales de aminoácidos con capacidad de coordinar los átomos metálicos. Puesto que los átomos metálicos contienen un poder de dispersión mayor que los que componen la proteína, éstos pueden llegar a modificar sensiblemente la intensidad del espectro de difracción del cristal derivativo frente al nativo, y estas diferencias de intensidad entre ambos espectros se pueden usar para calcular las posiciones de los átomos pesados en la celda unidad, lo cual nos permitirá obtener un conjunto de fases iniciales, que combinadas con los módulos de los factores estructura observados en el experimento, nos permitirá calcular un primer mapa de densidad electrónica (Anexo I).

Los pasos a seguir en un experimento de MIR son:

- Preparación de uno, o varios derivados de átomo pesado del cristal de la proteína. Un primer control del isomorfismo viene dado por la comparación de las dimensiones de la celda unidad del cristal nativo y del derivado.
- Colección de datos de difracción para los cristales de proteína nativa y los derivados.
- Determinación de las coordenadas de los átomos pesados mediante la función de Patterson.
- Refinamiento de los parámetros de los átomos pesados y cálculo de los ángulos de fase de la proteína.
- Cálculo de la densidad electrónica de la proteína.

En el presente trabajo se colectaron un buen número de cristales “derivativos” usando diferentes átomos pesados, (Tabla 6).

En un principio se colectaron varios cristales nativos, uno de los cuales difractó a 2.6 Å en un difractómetro convencional. Con el fin de obtener mejor resolución se colectaron cristales con difracción de sincrotrón y se obtuvieron dos cristales que difractaron a 2.3 y 2.5 Å (Tabla 4: 1,2). El cristal X3.2 (Tabla 4: 2), se colectó primero y se utilizó como nativo para los experimentos de MIR que llevamos a cabo a continuación.

Los primeros cristales derivativos que se colectaron se hicieron añadiendo BrNa (Tabla 6: 1-4) o  $K_2Pt(SCN)_6$  (Tabla 6: 13,14) a la solución de cristalización. Se

calcularon mapas de Patterson para varios cristales de cada tipo y se determinaron las posiciones (x, y, z) en la celda (espacio real) a partir de los picos en secciones Harker (x,z) y no-Harker (y), usando un átomo de Br como referencia (y=0.5). Posteriormente las posiciones de los átomos pesados se refinaron con MLPHARE (Collaborative Computational Project 1994) (Tabla 5). Como comprobación adicional se usaron las fases obtenidas con los sitios del bromo para encontrar los sitios del platino. Para ello se calcularon mapas de Fourier con las fases obtenidas a partir de los dos átomos de bromo, en los que se identificaron los átomos de Pt calculados en los mapas de Patterson.

Debido a la baja calidad de las fases se buscaron sitios adicionales en otros derivados: INa (Tabla 6: 24), K<sub>2</sub>AuCl<sub>4</sub> (Tabla 6: 25) y K<sub>2</sub>Pt(CN)<sub>4</sub> (Tabla 6: 6,7). En los cristales de I y de Au se obtuvieron varios sitios que no refinaron. Mientras que en los cristales que contenían K<sub>2</sub>Pt(CN)<sub>4</sub> se encontró un nuevo sitio, que mejoró las fases de obtenidas con anterioridad (Tabla 5).

	X	Y	Z	FACTOR-B
<i>K<sub>2</sub>Pt(SCN)<sub>6</sub></i>				
<b>1</b>	0.047	0.500	0.201	25.71
<b>2</b>	0.054	0.475	0.205	29.68
<i>PhP<sub>a</sub> = 1.46</i>				
<i>BrNa</i>				
<b>1</b>	0.049	0.490	0.202	23.37
<b>2</b>	0.131	0.259	0.059	85.8
<b>3</b>	0.776	0.023	0.521	64.06
<i>PhP<sub>a</sub> = 1.59</i>				
<i>K<sub>2</sub>Pt(CN)<sub>4</sub></i>				
<b>1</b>	0.951	-0.010	0.798	14.24
<i>PhP<sub>a</sub> = 1.51</i>				
<i>FOM = 0.4513</i>				

**Tabla 5 .** Posiciones y refinamiento de átomos pesados. Se indican las posiciones de los primeros sitios encontrados para los metales pesados K<sub>2</sub>Pt(SCN)<sub>6</sub>, BrNa y K<sub>2</sub>Pt(CN)<sub>4</sub>, así como sus estadísticas tras el refinamiento de las posiciones.

Las estadísticas obtenidas eran similares con los programas MLPHARE (Collaborative Computational Project 1994) y SHARP. Los datos se sometieron a “modificación de la densidad” utilizando el programa DM (Collaborative Computational Project 1994) con 60 y 40% de solvente, obteniéndose mejores estadísticas con alto contenido de solvente, lo que corroboraba la presencia de una única molécula en la unidad asimétrica del cristal. Con las fases obtenidas se calcularon unos mapas de densidad electrónica iniciales, pero eran de muy baja calidad, siendo imposible comenzar la reconstrucción la molécula.

DERIVATIVOS								
CRISTAL	COMPUESTO	LINEA	RESOLUCION	COMPL. (%)	REDUN.	I/ $\sigma$	R(%)	GRUPO ESPACIAL
1 Br10	BrNa	BM16	20-2.6	97.1	6	9 (2.2)	5.9 (33.5)	P21
2 Br100.pk	BrNa	BM16	20-2.6	97.9	5.9	5.9 (2.2)	6.9 (33.8)	P21
Br100.re	BrNa	BM16	20-2.6	99.1	4.4	6.9 (2.2)	6.1 (34.0)	P21
Br100.ip	BrNa	BM16	20-2.6	97.9	4.10	6.5 (2.0)	6.2 (3.6)	P21
3 Br x2	BrNa	ID14.4	25-3.2	99.5	5.2	4.4 (2.5)	12.0 (29.4)	P21
4 Br x3	BrNa	ID14.4	20-3.0	99.6	7.7	9.1 (3.0)	6.3 (24.4)	P21
5 Pt3x1	K <sub>2</sub> PtCl <sub>6</sub>	ID14.1	25-2.9	98.7	9.5	7.5 (2.7)	6.8 (27.2)	P21
6 Pt4	K <sub>2</sub> Pt(CN) <sub>4</sub>	BM16	25-3.3	99.8	4.8	9.1 (2.5)	7.9 (29.7)	P21
7 Pt4x1	K <sub>2</sub> Pt(CN) <sub>4</sub>	ID14.4	25-2.7	99.5	7.3	7.9 (1.8)	7.0 (37.7)	P21
8 Pt7x1	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID14.1	25-3.0	99.7	7.3	4.4 (1.8)	12.7 (40.9)	P21
9 Pt7x1 pk	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID23.1	25-3.2	99.7	7.3	8.6 (2.4)	7.6 (3.5)	P21
Pt7x1 ip	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID23.1	25-3.2	99.5	3.6	7.9 (1.9)	8.1 (36.6)	P21
Pt7x1 rm	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID23.1	25-3.2	99.5	3.6	7.2 (1.8)	8.5 (39.3)	P21
10 Pt7x2	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID23.1	25-3.2	99.6	7.4	6.1 (2.1)	8.5 (31.4)	P21
11 Pt7x4	K <sub>2</sub> Pt(NO <sub>2</sub> ) <sub>4</sub>	ID23.1	25-3.3	99.3	7.3	7.7 (1.9)	8.6 (36.6)	P21
12 Pt9	K <sub>2</sub> Pt(SCN) <sub>6</sub>	BM16	25-3.1	99.6	3.7	12.3 (2.7)	5.9 (27.9)	P21
13 Pt9 x1 pk	K <sub>2</sub> Pt(SCN) <sub>6</sub>	ID14.4	25-2.4	97.2	5.4	10 (1.9)	5.1 (38.4)	P21
Pt9 x1 re	K <sub>2</sub> Pt(SCN) <sub>6</sub>	ID14.4	25-2.6	98.4	3.4	13.4 (1.9)	4.3 (37.6)	P21
Pt9 x1 ip	K <sub>2</sub> Pt(SCN) <sub>6</sub>	ID14.4	25-2.5	98.8	5.5	10.6 (1.9)	5.0 (40.1)	P21
14 Pt9 x4	K <sub>2</sub> Pt(SCN) <sub>6</sub>	ID14.4	25-2.8	99.4	7.5	8.2 (1.8)	8.0 (42.0)	P21
15 Hg20x1	HgCl <sub>2</sub>	BM16	25-2.9	99.6	7.2	7.7 (1.9)	7.8 (37.7)	P21
16 Hg20x2	HgCl <sub>2</sub>	ID14.1	25-3.0	99.4	8.5	4.9 (1.8)	10.1 (42.2)	P21
17 Hg21	HgBr <sub>2</sub>	BM16	25-3.1	99.8	3.7	9.1 (2.2)	7.6 (33.1)	P21
18 Hg21x1	HgBr <sub>2</sub>	BM16	25-3.2	96.7	6.3	6.0 (2.0)	11.0 (35.8)	P21
19 Hg54x1	(C <sub>2</sub> H <sub>5</sub> Hg)Cl	BM16	25-4.0	98.6	5.4	2.5 (1.7)	25.6 (40.1)	P21
20 Hg54x2 pk	(C <sub>2</sub> H <sub>5</sub> Hg)Cl	BM16	25-3.0	99.8	7.1	7.7 (1.8)	8.5 (39.4)	P21
Hg54x2 rm	(C <sub>2</sub> H <sub>5</sub> Hg)Cl	BM16	25-3.0	99.6	7.1	8.4 (2.2)	7.3 (32.7)	P21
Hg54x2 ip	(C <sub>2</sub> H <sub>5</sub> Hg)Cl	BM16	25-3.0	99.6	7.0	7.4 (1.6)	8.0 (43.3)	P21
21 Gd55x1	C <sub>16</sub> H <sub>28</sub> GdN <sub>5</sub> O <sub>9</sub>	BM16	25-2.7	97.9	4.2	5.5 (2.4)	8.1 (31.3)	P21
22 Gd56x1	C <sub>17</sub> H <sub>29</sub> N <sub>4</sub> O <sub>7</sub> Gd	BM16	25-3.0	99.8	3.7	11.3 (1.9)	6.4 (40.1)	P21
23 Gd56	C <sub>17</sub> H <sub>29</sub> N <sub>4</sub> O <sub>7</sub> Gd	BM16	25-3.7	98.7	2.5	5.9 (2.4)	11.5 (31.1)	P21
24 I	INa	ID14.4	25-2.6	99.7	7.5	9.8 (2.0)	6.0 (37.6)	P21
25 Au	K <sub>2</sub> AuCl <sub>4</sub>	ID14.4	25-2.7	98.6	7.3	9.5 (1.8)	6.5 (41.8)	P21
26 Yb67x1	YbCl <sub>3</sub>	BM16	25-3.1	99.8	3.7	9.3 (2.2)	7.9 (35.2)	P21
27 Yb67x2	YbCl <sub>3</sub>	BM16	25-2.6	88.3	7.5	4.10 (2.0)	9.7 (37.6)	P21
28 Os35x1	K <sub>2</sub> O <sub>5</sub> Cl <sub>6</sub>	BM16	25-2.9	99.6	3.7	11.2 (2.1)	6.3 (37.5)	P21
29 Os35x2	K <sub>2</sub> O <sub>5</sub> Cl <sub>6</sub>	BM16	25-3.2	99.8	7.4	7.8 (2.1)	9.1 (36.0)	P21
30 Ir63x1	IrCl <sub>3</sub> · 3H <sub>2</sub> O	BM16	25-2.7	99.3	3.7	9.2 (2.2)	6.7 (35.2)	P21
31 Ir64x2	Na <sub>3</sub> IrCl <sub>6</sub>	BM16	25-3	99.0	3.8	9.9 (1.9)	7.4 (39.6)	P21
32 Lu68x1	LuCl <sub>3</sub> · 6H <sub>2</sub> O	BM16	25-2.6	99.8	3.7	7.0 (2.2)	6.8 (34.5)	P21
33 Lu69x1	Lu(OOCCH <sub>3</sub> ) <sub>3</sub>	BM16	25-2.7	99.2	3.7	9.7 (2.2)	5.7 (35.4)	P21
34 Pd23x3	K <sub>2</sub> PdCl <sub>6</sub>	BM16	25-2.7	99.4	3.7	7.6 (2.1)	6.8 (36.1)	P21
35 Pd24x1	K <sub>2</sub> PdCl <sub>4</sub>	BM16	25-3.0	99.8	8.3	7.4 (1.9)	9.1 (39.6)	P21
36 Co41x1	Na <sub>3</sub> (Co(NO <sub>2</sub> ) <sub>6</sub> )	BM16	25-2.7	98.6	3.8	12.5 (2.3)	4.10 (33.7)	P21
37 Ho70x1	HoCl <sub>3</sub> · 6H <sub>2</sub> O	BM16	25-3.0	99.0	7.5	7.8 (2.1)	7.8 (32.2)	P21
38 Sm33x1	SmCl <sub>3</sub> · 6H <sub>2</sub> O	ID23.1	25-3.0	99.5	6.2	7.3 (2.2)	8.2 (30.4)	P21
39 Pb17x1	Pb(CH <sub>3</sub> COO) <sub>2</sub> · 3H <sub>2</sub> O	BM16	25-2.8	99.8	3.7	10.2 (1.9)	6.7 (40.0)	P21
40 Ti66x1	TiCl <sub>3</sub>	BM16	25-2.8	99.3	3.7	8.9 (1.9)	7.7 (35.4)	P21
41 W49x1	NaW · 2H <sub>2</sub> O	BM16	25-3.0	99.1	3.7	8.6 (2.1)	7.8 (36.9)	P21

**Tabla 6.** Estadísticas de los datos de difracción colectados con cristales derivados de la formas cristalina X3. Los datos se procesaron con XDS (Kabsch 1993) y se escalaron con SCALA (Collaborative Computational Project 1994). Se indica la línea del ESRF donde fueron colectados, el compuesto utilizado para derivatizar el cristal, la resolución, la completitud, la redundancia, el valor de I sobre sigma, el factor R y el grupo espacial.

Para intentar mejorar las fases se prepararon cristales derivados con numerosos metales pesados y se colectaron datos cristalográficos en la línea BM16 del ESRF (Grenoble, Francia), (Tabla 6: 26-41). Si bien se identificaron varios sitios de átomos pesados en diversos derivados, sólo nuevos sitios caracterizados en cristales con  $K_2Pt(NO_2)_4$  dieron estadísticas relevantes. Asimismo, los mapas de densidad electrónica calculados a partir de los sitios de este derivado de platino junto con átomos de Br (Tabla 7) presentaron una mayor calidad, pudiéndose localizar las moléculas de IC5-4D.

	X	Y	Z	FACTOR-B
<i>K<sub>2</sub>Pt(CN)<sub>4</sub></i>				
<b>1</b>	0.433	0.500	0.322	27.02
<b>2</b>	0.924	0.401	0.903	31.25
<b>3</b>	0.391	0.476	0.29	66.27
<i>PhP<sub>a</sub> = 1.58</i>				
<i>BrNa</i>				
<b>1</b>	0.790	0.010	0.544	15.98
<b>2</b>	0.168	0.375	0.432	45.32
<b>3</b>	0.828	0.418	0.837	25.44
<b>4</b>	0.567	0.025	0.677	20.13
<i>PhP<sub>a</sub> = 1.00</i>				
<i>FOM = 0.3760</i>				

**Tabla 7.** Posiciones y refinamiento de átomos pesados. Se indican las posiciones de los sitios finales encontrados para los metales pesados  $K_2Pt(CN)_4$  y  $BrNa$ , así como sus estadísticas tras el refinamiento. Estos fueron los metales utilizados para calcular las fases iniciales.

## MAD

Para intentar mejorar la fases se llevaron a cabo varios experimentos de MAD con diferentes átomos pesados (Tabla 6: 2, 9, 13, 20), este método tiene la ventaja de todos los datos del experimento se colectan a partir de un solo cristal, evitando el problema de la falta de isomorfismo (Anexo I). Sin embargo, en ningún caso obtuvimos una señal anómala clara, por lo que no pudimos identificar sitios de unión de átomos pesados ni calcular fases cristalográficas.

## Reemplazo molecular

Inicialmente se usó el método de reemplazo molecular usando el programa AMORE y las estructuras cristalográficas de los dominios 1 y 2 de ICAM-2 (Casasnovas JM 1997) e ICAM-1 (Casasnovas JM 1998) respectivamente. No obstante no se obtuvieron soluciones claras, posiblemente por tratarse exclusivamente del 50% de la estructura. A partir de la publicación de la estructura de los dominios 3 y 4 de ICAM-1 (Yang Y 2004) se continuó con el trabajo de reemplazo molecular. Sin embargo, no se obtuvo una solución convincente hasta el uso del programa PHASER y buscando dominios individualmente. Usando esta estrategia se obtuvo una solución en la que los 4 dominios estaban concatenados correctamente (D1-D2-D3-D4), con los extremos N y C terminales a una distancia esperada. Los dominios se sometieron a refinamiento de cuerpo rígido con el programa REFMAC (Collaborative Computational Project 1994) y se calcularon unas fases iniciales.

Para comprobar la veracidad de la solución de reemplazo molecular y de las posiciones de los átomos pesados determinadas por el método MIR se calcularon mapas de diferencia de Fourier ( $F_{PH}-F_P$ ) con las fases provenientes del reemplazo molecular. Es estos mapas se pudieron identificar los sitios de Pt y Br identificados por métodos de Patterson (Tabla 7). En este caso las coordenadas (x,z) eran idénticas, mientras que la posición de la coordenada (y) estaba desplazada 0.375 en todos los sitios debido a una diferencia en el origen de los mapas calculados por las diferentes metodologías.

Las fases obtenidas de la solución de reemplazo molecular y las provenientes de los átomos pesados se combinaron usando el programa SIGMA-A (Collaborative Computational Project 1994). Los datos se sometieron a “*modificación de la densidad*” con DM (Collaborative Computational Project 1994) y se construyeron mapas de densidad electrónica con FFT (Collaborative Computational Project 1994), que presentaron mejor calidad que aquellos calculados con las fases de los átomos pesados o del reemplazo molecular.

## Reconstrucción y refinamiento de la estructura

Se inició la construcción del modelo de IC5-4D con mapas de densidad electrónica a 3 Å de resolución, manualmente con el programa “O” (Jones TA 1991). Una vez

encajada la cadena principal en los mapas de densidad electrónica se procedió a la sustitución de los residuos de los modelos por los de ICAM-5 y al modelado de las cadenas laterales. Algunos residuos de carbohidratos (NAG) unidos a Asn fueron fácilmente modelados en el mapa inicial, así como las conexiones entre dominios, donde algunos residuos fueron omitidos en los modelos. El primer modelo se refinó con el programa CNS (Brünger AT 1998) y datos cristalográficos cuya resolución se extiende a 2.3Å. El modelo se sometió a varios ciclos de minimización por “*simulated annealing*”, refinando el factor de temperatura (B-group) usando datos cristalográficos entre 15 y 3Å de resolución. Se calcularon mapas ( $F_o - F_c$  y  $2F_o - F_c$ ) con los que se procedió a una reconstrucción manual más precisa del modelo, añadiéndose algunos residuos de carbohidratos y regiones de bucles ausentes en el modelo inicial. Posteriormente el modelo se refinó con el programa CNS con varios ciclos de “*simulated annealing*”, minimización de la energía y refinamiento del factor de temperatura (B-group). El modelo refinó hasta valores de  $R_{libre}$  de aproximadamente 30% usando datos a 3Å de resolución.

La estructura actual (Fig.4.19B) muestra los cuatro dominios inmunoglobulina concatenados y algunos residuos de las N-glicosilaciones (Fig.4.19B rojo).

El residuo Glu<sup>37</sup>, fundamental para la unión a la integrina LFA-1, se encuentra situado al final de la hebra C del primer dominio, en una superficie plana formada por las hebras C y F (Fig.4.19B). Este motivo está muy conservado entre los miembros de la subfamilia (Fig.4.19), como por ejemplo el residuo Glu<sup>34</sup> de ICAM-1 (Fig4.19A). El fragmento comprendido entre las hebras B y E del dominio D4 presenta alta flexibilidad y no se encuentra claramente definido en el mapa de densidad electrónica. Esta zona está ausente en la molécula de ICAM-1 utilizada para hacer el modelo (Yang Y 2004), por su elevada flexibilidad.

#### **4.3.5. Determinación de la estructuras de la formas X1 y X2**

Una vez avanzado el trabajo de resolución de la estructura de la forma cristalina X3 de IC5-4D y mientras se procedía a su refinamiento se intentó determinar la estructura de las formas cristalinas que difractaban a baja resolución. Se utilizó el programa PHASER y dos fragmentos de la molécula que comprendían los dominios 1 y 2 y los dominios 3 y 4. Se obtuvieron claras soluciones para las dos formas cristalinas en la



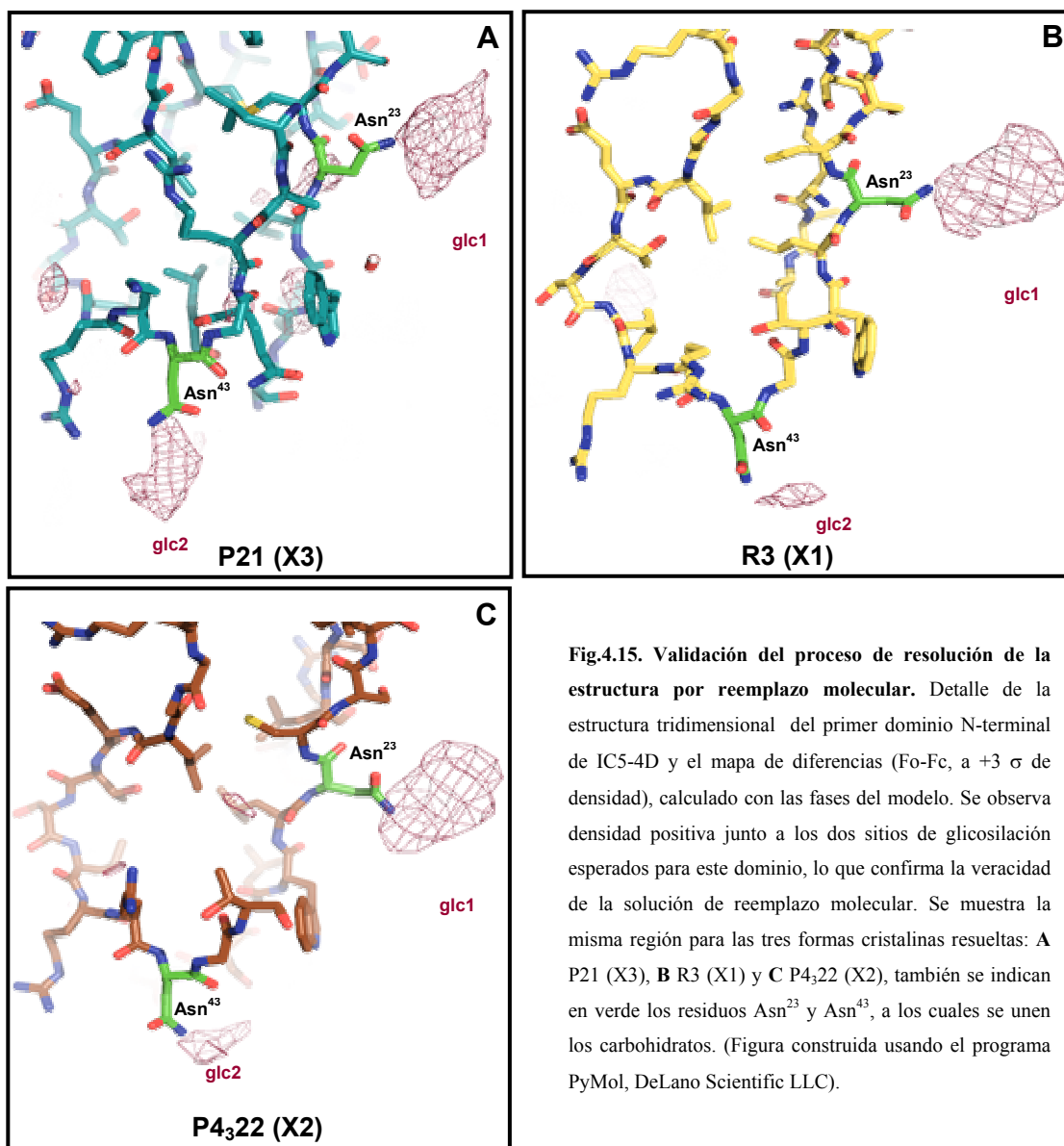
que los dos módulos de la proteína aparecían concatenados. El proceso de reemplazo molecular determinó que el correcto grupo espacial para la forma X2 era  $P4_322$  en vez del  $P4_122$ . Las soluciones se sometieron a refinamiento (cuerpo rígido y “restrain refinement”) con el programa REFMAC (Tabla 8).

	<i>X1 (H3)</i>	<i>X2 (P4<sub>3</sub>22)</i>
<i>Resolución (Å)</i>	15-4.2	15-3.9
<i>R<sub>factor</sub></i>	0.27873	0.31737
<i>R<sub>libre</sub></i>	0.31384	0.34966
<i>Desviación Rms enlaces (Å)</i>	0.11	0.008
<i>Desviación Rms ángulos (grados)</i>	1.392	1.247

**Tabla 8.** Estadísticas del refinamiento de las formas cristalinas X1 y X2 a baja resolución.

#### 4.3.6. Validación del proceso de resolución de las estructuras

Una prueba adicional de que las soluciones encontradas eran correctas la obtuvimos del hecho de que los mapas de densidad electrónica de diferencias (Fo-Fc) mostraron



**Fig.4.15. Validación del proceso de resolución de la estructura por reemplazo molecular.** Detalle de la estructura tridimensional del primer dominio N-terminal de IC5-4D y el mapa de diferencias (Fo-Fc, a +3  $\sigma$  de densidad), calculado con las fases del modelo. Se observa densidad positiva junto a los dos sitios de glicosilación esperados para este dominio, lo que confirma la veracidad de la solución de reemplazo molecular. Se muestra la misma región para las tres formas cristalinas resueltas: **A** P21 (X3), **B** R3 (X1) y **C** P4<sub>3</sub>22 (X2), también se indican en verde los residuos Asn<sup>23</sup> y Asn<sup>43</sup>, a los cuales se unen los carbohidratos. (Figura construida usando el programa PyMol, DeLano Scientific LLC).

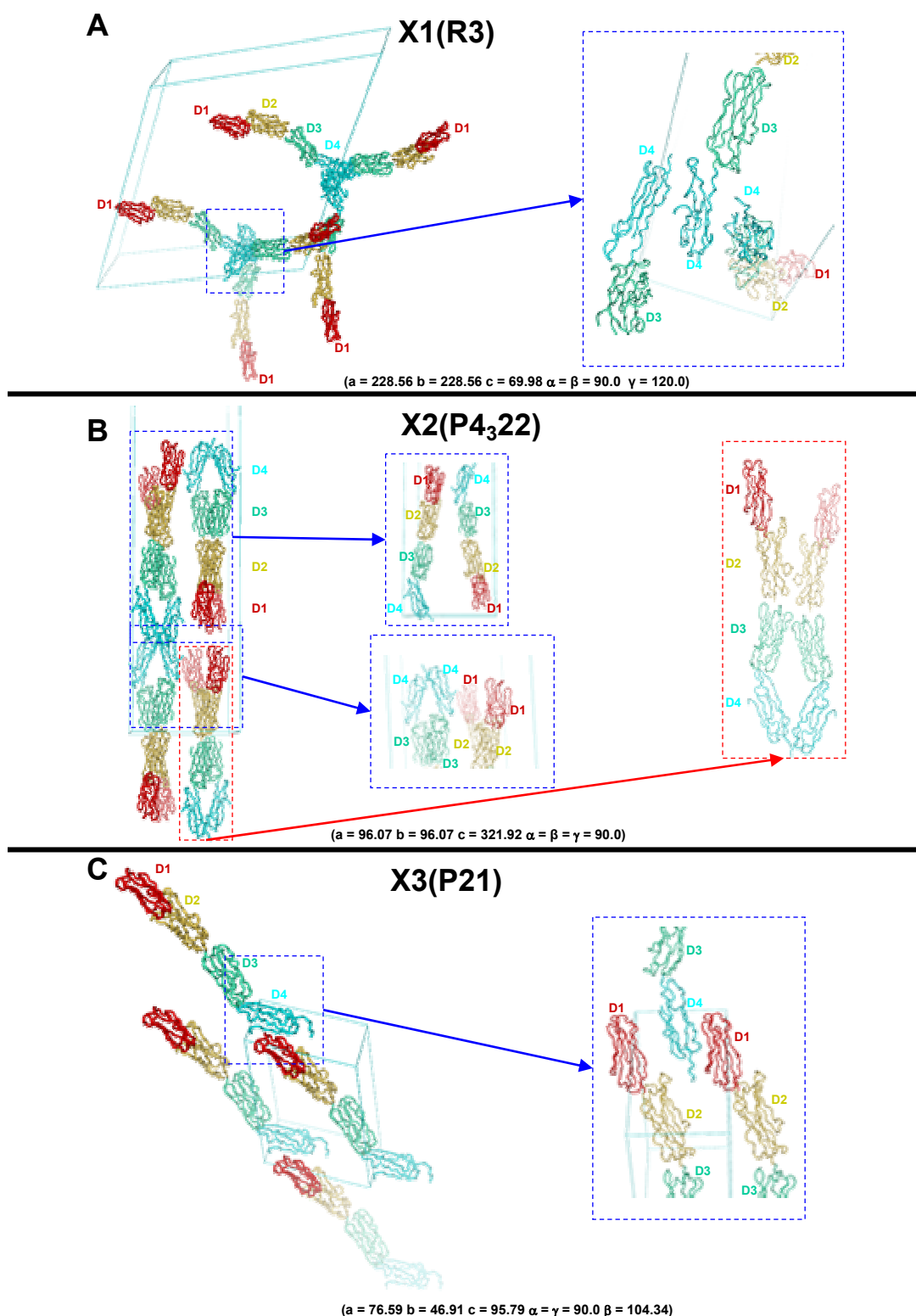
densidad positiva adyacente a los residuos Asn glicosilados (Fig.4.15). Teniendo en cuenta que los modelos utilizados para la resolución de las estructuras carecían de glicosilaciones, la presencia de densidad corrobora la veracidad de las soluciones obtenidas por reemplazo molecular para las tres formas cristalinas.

#### 4.3.7. Empaquetamiento de las moléculas dentro de la celda unidad en las tres formas cristalinas resueltas

A pesar de tener la misma proteína en las tres formas cristalinas resueltas, (X1, X2 y X3), el empaquetamiento de las moléculas era claramente diferente entre ellas (Fig.4.16). Pensamos que las interacciones intermoleculares observadas en los cristales pueden definir la interacción de adhesión celular homofílica característica de ICAM-5.

La forma cristalina X1 pertenece al grupo espacial R3 y contiene una sola molécula en la unidad asimétrica. Su celda unidad es relativamente grande ( $a = 228.56$   $b = 228.56$   $c = 69.98$   $\alpha = \beta = 90.0$   $\gamma = 120.0$ ). Se observa como tres moléculas relacionadas por simetría interaccionan a través del dominio 4 (Fig.4.16A).

La forma cristalina X2 pertenece al grupo espacial P4<sub>3</sub>22, ( $a = 96.07$   $b = 96.07$   $c = 321.92$   $\alpha = \beta = \gamma = 90.0$ ) y al contrario que la anterior contiene dos moléculas por unidad asimétrica. Estas dos moléculas se empaquetan formando un homodímero en *cis*, en el que se observan interacciones entre los dominios 2, 3 y 4 (Fig.4.16B recuadrado en rojo). A su vez, los homodímeros interaccionan en *trans* con otros dímeros adyacentes dispuestos en forma antiparalela a través de los dominios D1 N-terminal y D4 C-terminal (Fig.4.16B recuadrado en azul). Asimismo, se observan contactos cristalinos entre la región superior (D1-D1) e inferior de homodímeros (D4-D4) adyacentes.



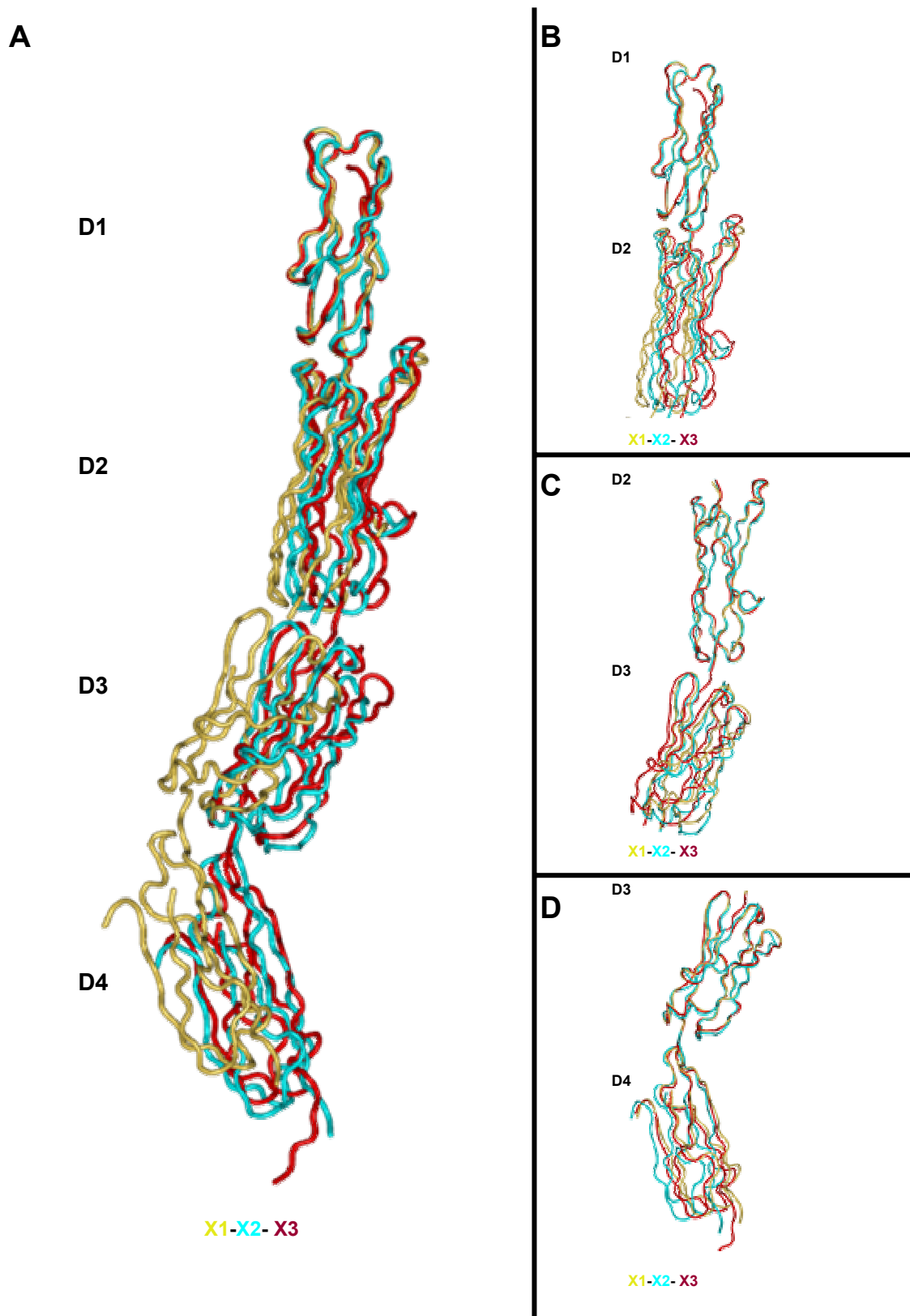
**Fig.4.16. Empaquetamiento de las moléculas en la celda unidad.** Las tres formas cristalinas resueltas se empaquetan de forma diferente en la celda unidad. **A. Forma cristalina X1**, pertenece al grupo espacial R3 y tiene una molécula por unidad asimétrica. Recuadrado en azul se muestra un detalle de la zona donde convergen las moléculas tres a tres con sus dominios D4 muy próximos entre sí. **B. Forma cristalina X2**, pertenece al grupo espacial P4<sub>3</sub>22 y contiene dos moléculas por unidad asimétrica, que forman un dímero en *cis* (recuadrado en rojo), mediante interacciones entre los dominios 2, 3 y 4. Recuadrado en azul se muestran las posibles interacciones entre las moléculas en *trans* de dos dímeros adyacentes. **C. Forma cristalina X3**, pertenece al grupo espacial P21 y tiene una molécula en la unidad asimétrica. Las moléculas se ordenan dentro de la celda unidad adoptando una configuración cola-cabeza, en el recuadro azul se detalla como dos moléculas orientadas con el dominio D1 hacia arriba se empaquetan muy cerca del dominio D4 de otra molécula adyacente. Para las tres estructuras se muestra el dominio 1 en rojo, el dominio 2 en amarillo, el dominio 3 en verde y el dominio 4 en azul. (Las imágenes se prepararon usando el programa PyMol, DeLano Scientific LLC).

Por último analizamos la forma cristalina X3, de la cual obtuvimos la estructura a alta resolución y es con la que más hemos trabajado. Pertenece al grupo espacial P21 y las proporciones de su celda son más pequeñas que en las anteriores ( $a = 76.59$   $b = 46.91$   $c = 95.79$   $\alpha = \gamma = 90.0$   $\beta = 104.34$ ). Contiene una sola molécula por unidad asimétrica y eso hace que el factor de temperatura (B) de los átomos sea elevado, lo cual dificultó su refinamiento. Las moléculas se disponen entre ellas de forma paralela en disposición cola-cabeza, donde el dominio D1 de una molécula se sitúa muy cerca del dominio D4 de otra molécula adyacente (Fig.4.16C). Asimismo se observan interacciones entre los dominios D1 y D2 con el dominio C-terminal D4.

#### **4.3.8. Estudio de la movilidad entre los dominios en las tres formas cristalinas**

A continuación se pasó a analizar la movilidad entre los dominios de las tres formas cristalinas resueltas. Para ello se alinearon las tres moléculas completas superponiendo los dominios D1 de cada una de ellas (Fig.4.17A). Este alineamiento muestra que las estructuras de X2 y X3 son más parecidas entre si que la de X1, así como cierta flexibilidad entre dominios.

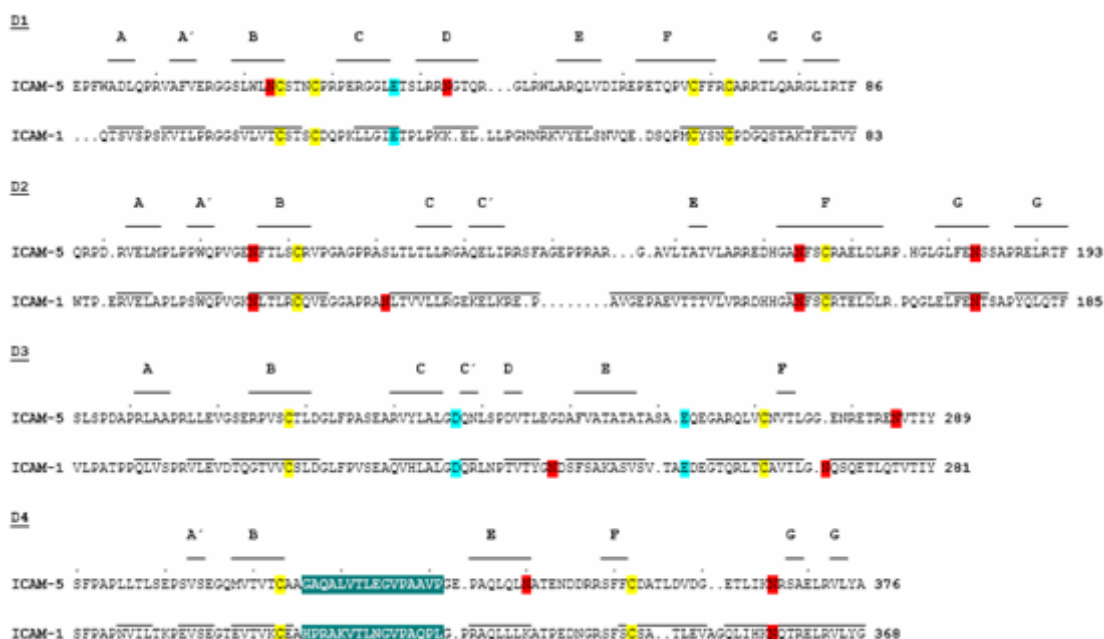
Para determinar donde residían las principales diferencias en la disposición de los dominios se alinearon de dos en dos superponiendo uno de ellos (Fig.4.17B, C y D). En la unión entre los dominios 1 y 2 es donde se aprecia más diferencia, especialmente entre la forma cristalina X1 y las otras dos (Fig.4.17B). Esta diferencia en la zona de unión de los dos primeros dominios determina que la forma cristalina X1 tenga una conformación diferente a las otras dos (Fig.4.17A). El ángulo entre los dominios 2 y 3 es prácticamente igual en las estructuras X1 y X2 las cuales difieren un poco de la forma X3 (Fig.4.17B). Los dominios 3 y 4 mantienen una conformación similar entre las formas X1 y X3, mientras que la conformación de X2 (P4<sub>3</sub>22) cambia ligeramente.



**Fig.4.17. Movimiento entre los dominios y comparación de los mismos en las tres formas cristalinas.** Se muestran los esqueletos de los Ca de las moléculas de ICAM-5 en las tres formas cristalinas resueltas: X1 (R3) en amarillo, X2 (P4322) en azul y X3 (P21) en rojo. **A.** Se muestran los dominios 1 y 2 tras la superposición del dominio 1. **B.** Dominios 2 y 3 tras la superposición del dominio 2. **C.** Dominios 3 y 4 tras la superposición del dominio 3.

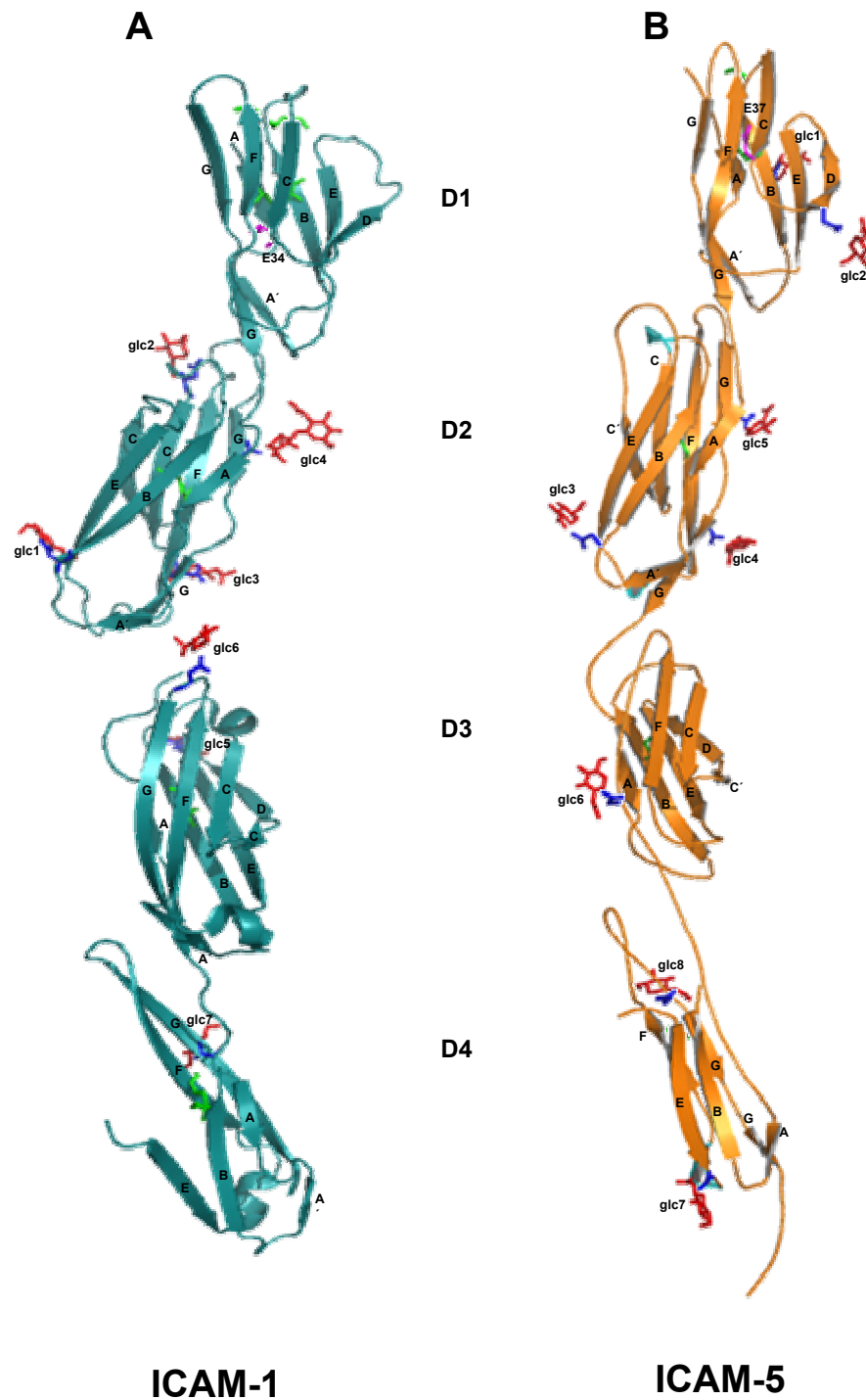
### 4.3.9. Análisis comparativo de la estructura de ICAM-5 con otros miembros de la subfamilia de los ICAMs

Como se ha dicho anteriormente, ICAM-5 es una glicoproteína de membrana que pertenece a la familia de las inmunoglobulinas y tiene entre un 38 y un 55% de identidad en la secuencia con los otros miembros de la subfamilia. Por lo tanto era de esperar que la estructura de ICAM-5 (Fig.4.19B) fuera muy parecida a las ya resueltas para ICAM-1 e ICAM-2 (Fig.1.7).



**Fig.4.18. Alineamiento estructural de los cuatro primeros dominios N-terminal de ICAM-5 con ICAM-1.** Los residuos cisteína, muy conservados están resaltados en amarillo. Los sitios de glicosilación están marcados en rojo y los fundamentales para la unión a integrinas en azul claro. Los 16 residuos desordenados en el cuarto dominio de ICAM-5 se indican en verde. El principio de cada dominio está indicado, y cada 10 residuos se marca con un punto encima de la secuencia de IC5, las hebras  $\beta$  se indican encima de la secuencia de ICAM-5. (El alineamiento se realizó utilizando el programa UCSF Chimera).

La identidad de secuencia entre los cuatro dominios N-terminal de ICAM-1 y de ICAM-5 es del 49.2%. Estructuralmente se observa que todos los puentes disulfuro permanecen iguales (Fig.4.18, Fig.4.19B). El residuo de unión a la integrina LFA-1 ( $\alpha_L\beta_2$ ), Glu<sup>37</sup> está conservado en todos los miembros de la familia, también lo están los residuos Asp<sup>237</sup> y Glu<sup>262</sup> (Fig.4.19B) que son fundamentales en la unión de ICAM-1 con la integrina Mac-1 ( $\alpha_M\beta_2$ ), a pesar de que no se ha descrito que esta integrina se una a IC5.



**Fig.4.19. Estructura cristalina de los cuatro dominios N-terminal de ICAM-1 e ICAM-5. A. Modelo ICAM-1 D1-D4,** fue construido uniendo las estructuras ya resueltas de ICAM-1 D1-D2 (Casasnovas JM 1998) e ICAM-1 D3-D5 (Yang Y 2004). Las láminas- $\beta$  se muestran en azul. Las N-glicosilaciones se indican en rojo, el residuo Glu<sup>34</sup> unido a la hebra C del primer dominio en morado y los puentes disulfuro en verde. **B. Estructura resuelta de ICAM-5 D1-D4.** Las láminas- $\beta$  se muestran en naranja y las  $\alpha$ -hélices en azul. Las N-glicosilaciones se indican en rojo, el residuo Glu<sup>37</sup> unido a la hebra C del primer dominio en morado y los puentes disulfuro en verde. El fragmento comprendido entre los residuos Gly<sup>316</sup> y Pro<sup>331</sup> entre las hebras B y E del cuarto dominio no se pudo construir. (Las imágenes se prepararon usando el programa PyMol, DeLano Scientific LLC).



El dominio 1 de ICAM-5 comprende los residuos entre Glu<sup>1</sup> y Phe<sup>86</sup> (Fig.4.18), pertenece al grupo de dominios inmunoglobulina de tipo I1 (Harpaz Y 1994), estos dominios se caracterizan por tener dos láminas  $\beta$  enfrentadas, una de las cuales contiene las hebras ABED mientras que la otra tendrá las A'GFC (Fig.4.19B). Este dominio es muy similar estructuralmente al de ICAM-1, la principal diferencia radica en la orientación de las lámina  $\beta$  D y E (Fig.4.19B). El residuo Glu<sup>37</sup>, fundamental para la unión de LFA-1, es el último residuo de la hebra C (Fig.4.19B). Presenta dos N-glicosilaciones unidas a los residuos Asn<sup>23</sup> y Asn<sup>43</sup>, visibles en la estructura (Fig.4.19B). Esta disposición de los azúcares es similar al del primer dominio de ICAM-2 (Fig.1.7A, C), mientras que ICAM-1 carece de glicosilaciones en este dominio (Fig.1.7B, D). Este dominio con 15 argininas en ICAM-5 es altamente básico (pI = 11.59) (Fig.4.20C), por lo que puede estar implicado en los procesos de interacción homofílica que lleva a cabo la proteína (Tian L 2000). Curiosamente el dominio D1 de ICAM-1 (Fig.4.20B) es prácticamente neutro (pI = 7.82).

El dominio 2 está formado por los residuos comprendidos entre Gln<sup>87</sup> y Phe<sup>193</sup> (Fig.4.18) y es estructuralmente muy similar al dominio 2 de ICAM-1 (Fig.4.19B, Fig.4.20A), por lo que se puede clasificar en el tipo I2 (Casasnovas JM 1998). La mayor diferencia estructural entre los dominios de ICAM-1 y de ICAM-5 está en el bucle C'E. La distribución de las hebras en las láminas  $\beta$  del dominio es del tipo ABE-A'GFCC'. Presenta tres sitios de N-glicosilación en las posiciones Asn<sup>106</sup>, Asn<sup>164</sup> y Asn<sup>183</sup> (Fig.4.19B), los tres fueron identificados en el mapa de densidad electrónica. Los azúcares unidos a Asn<sup>106</sup> y Asn<sup>164</sup> se colocan en la parte de abajo del dominio muy cerca de la zona de unión con el dominio 3 (Fig.4.19B). El unido a Asn<sup>183</sup>, se sitúa en la parte superior del dominio, en la hebra G (Fig.4.19B). Se distribuyen de forma similar que en ICAM-1, salvo que ICAM-1 tiene un azúcar más en el residuo Asn<sup>116</sup> situado en la parte superior del giro BC muy próximo a la unión con el primer dominio (Fig.1.7B). Este dominio, al igual que el primero, también es básico, (pI = 11.26) (Fig.4.20C).

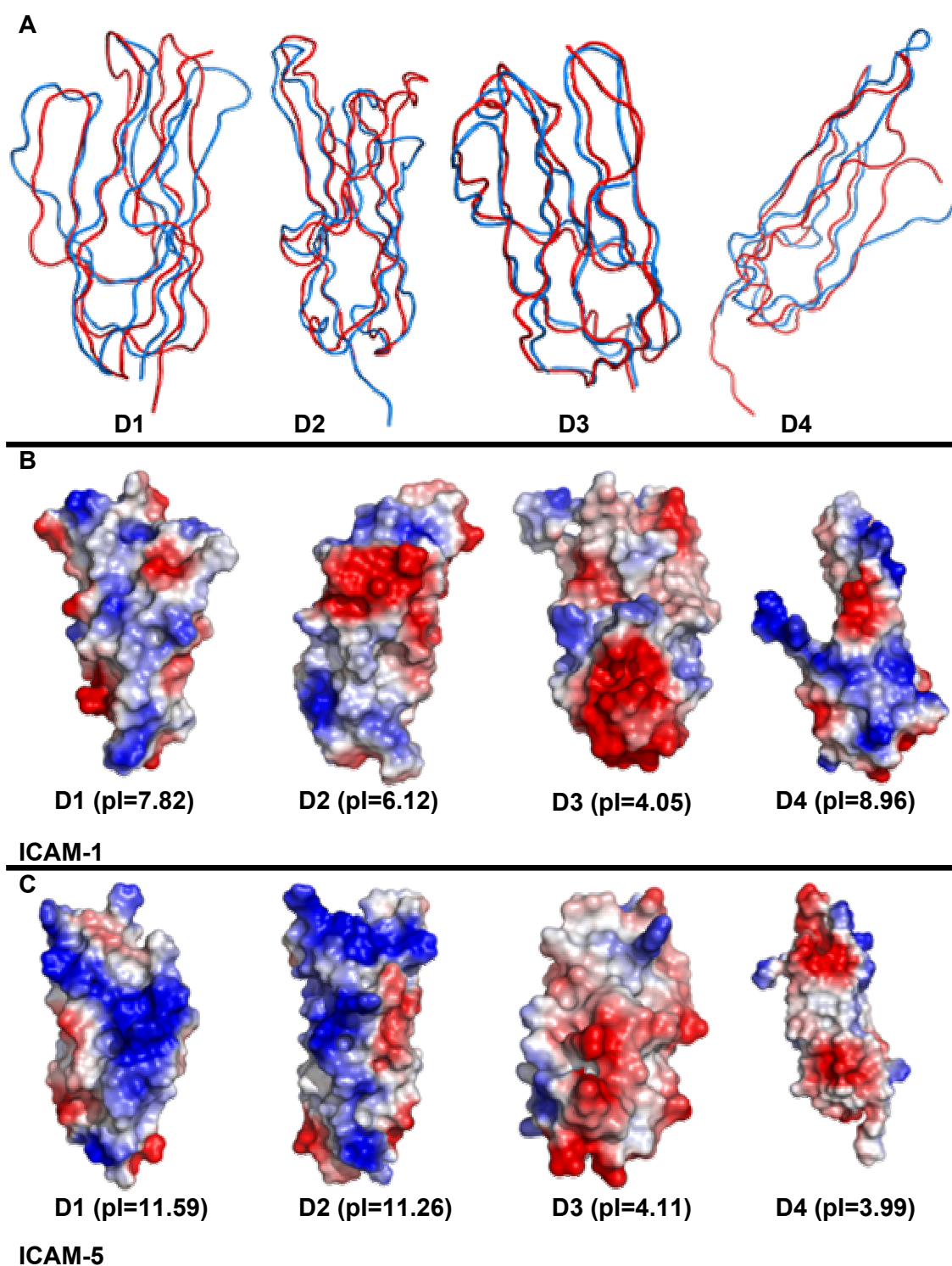
El dominio 3 está comprendido entre los residuos Ser<sup>194</sup> y Phe<sup>291</sup> (Fig.4.18), al igual que el anterior, es muy parecido al equivalente de ICAM-1 (Fig.4.19, Fig.4.20A). Se podría englobar dentro del grupo I1 de dominios inmunoglobulina. En este grupo las hebras se disponen según el esquema ABED-A'GFCC', aunque en nuestro modelo no se pudieron distinguir las hebras A' ni G. Posee los residuos conservados Asp<sup>237</sup> y Glu<sup>262</sup>, fundamentales para la unión a la integrina Mac-1 ( $\alpha_M\beta_2$ ), a pesar de no haberse descrito unión a dicha integrina para ICAM-5. Tiene tres posibles sitios de

glicosilación, (Asn<sup>239,272,285</sup>), aunque sólo se observó densidad significativa junto a Asn<sup>285</sup> (Fig.4.19B, Fig.4.20). Al contrario que ICAM-1, que contiene dos azúcares en la parte superior del dominio (Fig.1.7B), ICAM-5 no tiene ningún sitio de glicosilación en la zona interdominios entre D2 y D3 (Fig.4.19).

La molécula de ICAM5 resuelta tiene una forma claramente alargada, pero en la zona de unión entre los dominios D3 y D4 se dispone de forma oblicua (Fig.4.19B), adoptando una configuración parecida a ICAM-1 (Yang Y 2004).

El dominio 4 está formada por los residuos comprendidos entre Pro<sup>292</sup> y Ala<sup>376</sup> (Fig.4.19B), al igual que el resto de dominios está formado por dos láminas- $\beta$ , aunque en este caso solo se pudieron reconstruir 5 hebras en posición BE y A'GF, respectivamente. Se observaron glicosilación unidas a los residuos Asn<sup>340</sup> y Asn<sup>366</sup> (Fig.4.19B). La estructura del dominio es muy parecida a la de ICAM-1 (Yang Y 2004) (Fig.4.20A). El fragmento comprendido entre los residuos Gly<sup>316</sup> y Pro<sup>331</sup> presenta una alta flexibilidad, y no pudo ser reconstruido, puesto que los mapas no mostraban densidad electrónica en la zona equivalente a esta parte de la molécula, este fragmento también estaba ausente en la estructura de ICAM-1, que fue usada como modelo para resolver la estructura. Este dominio junto con el anterior son fuertemente ácidos (Fig.4.20C), con lo que pueden establecer interacciones homofílicas con los dos primeros dominios de otras moléculas adyacentes (Tian L 2000).

Curiosamente, aunque las moléculas de ICAM-1 e ICAM-5 son muy parecidas desde el punto de vista estructural, a nivel de su composición son sustancialmente distintas, esto hace que las condiciones electrostáticas de los dominios presenten diferencias significativas. Así en ICAM-1 los dominios D1 (pI = 7.82) y D2 (pI = 6.124) son prácticamente neutros, mientras que D3 (pI = 4.05) es fuertemente ácido y D4 (pI = 8.96) es básico (Fig.4.20B). Sin embargo en ICAM-5, los dominios D1 (pI = 11.59) y D2 (pI = 11.26) son claramente básicos, mientras que D3 (pI = 4.11) y D4 (pI = 3.99) son ácidos (Fig.4.20C). Esta distribución de cargas en ICAM-5 puede determinar su capacidad de mediar interacciones homofílicas.



**Fig.4.20. A. Comparación de las estructuras tridimensionales de ICAM-1 (Azul) e ICAM-5 (Rojo).** Alineación estructural de los diferentes dominios de ICAM-1 e ICAM-5 por separado. **B. Esquema de distribución de cargas en los cuatro primeros dominios N-terminal de ICAM-1.** Los dominios D1 y D2 son prácticamente neutros, el dominio D3 es ácido y el D4 básico. **C. Esquema de distribución de cargas en los cuatro primeros dominios N-terminal de ICAM-5.** Se puede ver como los dominios 1 y 2 son fuertemente básicos, mientras que los dominios 3 y 4 son fundamentalmente ácidos.



## **CAPITULO 5: DISCUSIÓN**



## 5. DISCUSION

### 5.1. Contribución de las N-glicosilaciones a las propiedades funcionales de los ICAMs

Los resultados presentados muestran como las N-glicosilaciones regulan funciones de la subfamilia de los receptores celulares denominados ICAMs.

El análisis de las estructuras cristalinas de ICAM-1 e ICAM-2 glicosilados (Casasnovas JM 1997; Casasnovas JM 1998) mostró que los azúcares más conservados (Fig.4.1) contactan con residuos de la proteína y tienen menos flexibilidad que el resto de azúcares, ya que presentan un factor de temperatura más bajo (Tabla 1), lo que indica que son críticos para la conformación de los dominios y consiguientemente para el reconocimiento de ligandos. Por lo tanto, parece ser que los azúcares que tienen los factores de temperatura más bajos pueden modular las estructuras tridimensionales de las glicoproteínas de células y virus. Los azúcares no conservados, como los del segundo dominio de ICAM-2 (Fig.4.1, Fig.4.10), no contribuyen de forma significativa a la estructura del receptor, pero están implicados en reconocimiento de ligandos en la familia de los ICAMs.

El alto grado de conservación del motivo azúcar-triptófano (Fig.4.3) en el centro del primer dominio N-terminal de la subfamilia de los ICAMs y su ausencia en ICAM-1 de humano y de chimpancé es bastante llamativa (Figs.4.1 y 4.3). Como se muestra aquí para ICAM-2 (Tabla 3), la supresión de este motivo afecta de forma importante a la unión de la integrina, a pesar de estar en la cara opuesta a la superficie de contacto. El análisis de unión a anticuerpos monoclonales mostró que el motivo azúcar-triptófano parece crítico para asegurar el correcto plegamiento del dominio inmunoglobulina y la conformación de la superficie de unión a la integrina, además podría proporcionar cierta estabilidad al dominio necesaria para mediar interacciones de adhesión celular. Se ha descrito un efecto conformacional similar para un carbohidrato en el caso de CD2, aunque en ese caso los azúcares contactan con un grupo de residuos cargados positivamente (Wyss DF 1995). La ausencia del sitio de glicosilación conservado glc1 en la parte central del primer dominio de ICAM-1 deja al descubierto las hebras- $\beta$  ABED del dominio para posibles interacciones moleculares. Esta superficie es usada en el proceso de dimerización del receptor y posibilita una interacción multivalente con la

integrina LFA-1 (Casasnovas JM 1998; Shimaoka M 2003), lo que aumenta la avidéz de la unión y proporciona un mecanismo de adhesión más especializado para la unión de leucocitos al endotelio (Springer 1994).

La falta de glicosilaciones en el dominio N-terminal del ICAM-1 humano (Fig.4.1) hace que presente una gran superficie muy propicia para interacciones proteína-proteína que no pueden darse en otros miembros de la subfamilia, lo cual ha influido positivamente para convertir a ICAM-1 en un receptor de patógenos. La presencia del primer azúcar conservado en ICAM-2 podría afectar al reconocimiento de proteínas provenientes de *P.falciparum*, porque estas se unen a residuos cercanos en la hebra- $\beta$  B de ICAM-1 (Casasnovas JM 1998). Además, mostramos que la inclusión del motivo azúcar-triptófano en ICAM-1 humano inhibe parcialmente la unión al principal grupo de rinovirus humano (Tabla 2), por lo que una molécula de ICAM-1 glicosilada en su dominio N-terminal funcionaría peor como receptor de rinovirus.

Con el propósito de entender la importancia de los carbohidratos en el plegamiento y la expresión de los dominios Ig en la subfamilia de los ICAMs, estudiamos el efecto producido por la introducción de N-glicosilaciones en el primer dominio de ICAM-1. Hasta el momento, no había sido posible expresar el primer dominio de ICAM-1 aislado, mientras que sí era posible la expresión del primer dominio de ICAM-3 aislado. La principal diferencia entre ambas moléculas radica en la falta de glicosilaciones en el caso de ICAM-1. Para determinar si las N-glicosilaciones son importantes para el plegamiento del dominio de unión a la integrina de los ICAMs, se introdujeron glicosilaciones en posiciones específicas del dominio de ICAM-1. La expresión de estos mutantes en moléculas de ICAM-1 con los dos dominios N-terminales se expresaban de forma similar al silvestre (Fig.4.4B). Sin embargo moléculas de ICAM-1 con un solo dominio silvestre o con azúcares introducidos en ciertas posiciones del inicio del dominio (hebras  $\beta$  B y E), como IC1-1D/T20N y Q58S, no se expresaban prácticamente nada. Mientras que los mutantes con glicosilaciones en la última hebra  $\beta$  del dominio (G), IC1-1D/Q73N, F79N, Y83N y W84N+P86T, presentaban unos niveles de expresión similares a los de las moléculas con dos dominios (Fig.4.4B y C). Estos resultados sugieren que los carbohidratos facilitan el plegamiento del dominio N-terminal de ICAM-1 aislado, que está implicado en el reconocimiento de la integrina LFA-1 y del rinovirus humano. Además se demuestra que la expresión de la proteína es dependiente de la posición donde se introduzca la glicosilación, por lo que la



localización del carbohidrato es crítica para el plegamiento del dominio de inmunoglobulina.

## **5.2. Importancia de las N-glicosilaciones en la unión de ICAMs a la lectina DC-SIGN**

Las estructuras cristalinas de ICAM-1 e ICAM-2 (Casasnovas JM 1997; Casasnovas JM 1998) localizaron las N-glicosilaciones en los dos dominios N-terminales de la subfamilia de los ICAMs (Fig.1.7). La mayoría de los azúcares en los ICAMs humanos se localizan en la parte superior o inferior de los dominios inmunoglobulina (Fig.1.7 y 4.1), cerca de las regiones interdominio. Esta localización preferente sugiere que los azúcares pueden influir en la orientación y flexibilidad de los dominios, así como en la localización de estas moléculas receptoras en la superficie celular. Un ejemplo de este comportamiento la constituye el azúcar unido a la región superior de la lámina- $\beta$  G del dominio 2 de ICAM-1 (glc4) (Fig.4.2), que participa en la región entre dominios interactuando con un gran número de residuos proteicos. La importante contribución de glc4 al plegamiento y conformación de los dos dominios N-terminal de ICAM-1 está demostrada por mutagénesis (Staunton DE 1990; Bella J 1998). Su conservación dentro de la subfamilia (Fig.4.1) sugiere que podrían constituir uniones interdominio en otros ICAMs relacionados.

Además de la importancia estructural de algunas N-glicosilaciones en la subfamilia de los ICAMs, hemos demostrado que éstas son usadas en el reconocimiento de ligandos, proporcionando especificidad en la unión de la lectina DC-SIGN. Para ello diseñamos un ensayo de unión proteína-proteína con el que detectamos unión específica de la lectina a ICAM-2 e ICAM-3 con sus dos dominios N-terminal y pérdida de unión a ICAM-1 (Fig.4.8), como ya había sido descrito anteriormente (Figdor CG 2002). Para demostrar que la unión de los ICAMs a la lectina era dependiente de azúcares se hicieron experimentos con la glicosidasa endo-H, mostrando que la actividad de unión de la lectina a los ICAMs está relacionada con la presencia de carbohidratos con alto contenido en manosas en el segundo dominio de ICAM-2 e ICAM-3 (Fig.4.8 y 4.10). Sin embargo, existe heterogeneidad en la composición de los azúcares de ICAM-2 e ICAM-3 y sólo alrededor del 40% de las moléculas tienen azúcares con alto contenido en manosas responsables de la unión a DC-SIGN (Fig.4.9, 4.8C). El porcentaje

calculado de moléculas de ICAM-2 que contienen oligosacáridos con alto contenido en manosa es similar al previamente descrito para ICAM-3 (Funatsu O 2001). Con estos experimentos se demostró que la unión es dependiente de azúcares y que no hay interacciones importantes dependientes de los residuos proteicos de los ICAMs con la lectina DC-SIGN.

Proteínas solubles con los dos primeros dominios N-terminal de ICAM-1 carecían de azúcares sensibles a endo-H y por consiguiente no se unían a la lectina DC-SIGN (Fig.4.8). Para demostrar que la unión de los ICAMs a la lectina era dependiente de los azúcares y no de la proteína, se construyeron mutantes de ICAM-1 en los cuales se introdujeron azúcares en los mismos sitios de glicosilación que tiene ICAM-2. La inclusión de un solo azúcar en tres posiciones diferentes del dominio 1 proporcionaba una población de moléculas con azúcares ricos en manosas, sugiriendo que a partir de un número umbral de glicosilaciones, que posiblemente depende de la proteína, se incorporan azúcares simples a los sitios de N-glicosilación. Sin embargo, la presencia de carbohidratos con alto contenido en manosa parece específica de los ICAMs, ya que este tipo de carbohidratos estaban ausentes en el receptor Nectin-1 con dos dominios de inmunoglobulina y un elevado número de carbohidratos de tipo N. Estas diferencias podrían estar relacionadas con diferencias en el plegamiento de los dominios inmunoglobulina entre las proteínas Nectin e ICAM así como con diferencias en los contactos interdominos.

Tras someter a electroforesis los mutantes de ICAM-2  $\Delta$ glc1,2,3 y  $\Delta$ glc4,5,6 se comprobó que el mutante que carece de glicosilaciones en el primer dominio es de menor tamaño que el que no tiene glicosilaciones en el segundo dominio (Fig.4.11). Lo que indica que los azúcares unidos al segundo dominio son más sencillos, como es el caso de los azúcares con alto contenido en manosa. Este hecho, sumado a la falta de funcionalidad del mutante  $\Delta$ glc4,5,6 y de moléculas de ICAM-3 con un solo dominio (Fig.4.8A), demuestra que los azúcares con manosa que reconoce la lectina se encuentran en el segundo dominio de los ICAMs.

Además de proporcionar protección contra proteasas y prevenir interacciones no específicas, los carbohidratos de tipo N regulan tanto la estructura como las funciones de los receptores, bien interaccionando con los residuos proteicos a los que están unidos, modulando la accesibilidad a las superficies de contacto proteicas o bien participando en el reconocimiento de ligandos. Como mostramos aquí con los ICAMs, la extensión y heterogeneidad de las glicosilaciones de tipo N influyen en el

plegamiento y en el reconocimiento de ligandos de esta familia de receptores celulares. Las glicosilaciones parecen estar íntimamente relacionadas con las propiedades funcionales de las glicoproteínas celulares y virales.

### **5.3. Estructura de ICAM-5 y modelos de interacción homofílica**

#### **5.3.1. Preparación de la proteína**

Para llevar a cabo los estudios estructurales de ICAM-5 se prepararon varias construcciones que codificaban para diferentes dominios funcionales (Fig. 4.12). Aunque finalmente nos centramos en la proteína que contenía los cuatro primeros dominios Ig (IC5-4D), ya que fue la primera proteína cristalizada.

La proteína se expresó en células CHO-Lec utilizando el método de la glutamina sintasa (Casasnovas JM 1998). Estas células introducen glicosilaciones con alto contenido en manosas, importantes para la expresión y el correcto plegamiento de la proteína. IC5-4D posee ocho sitios de N-glicosilación, (Fig.4.18). La proteína fue purificada manteniendo los azúcares, lo que mejoró la estabilidad y solubilidad de la proteína, además de mantenerla en un estado más próximo al fisiológico.

Debido a que los dominios extracelulares de ICAM-5 están implicados en procesos de adhesión homofílica (Tian L 2000), cabría pensar que las diferentes moléculas de la proteína pudieran interaccionar formando dímeros. Sin embargo uno de los pasos de purificación de la proteína consistía en una cromatografía de exclusión molecular utilizando una columna Superdex-200 (Amersham Biosciences) calibrada, que mostró que la proteína se comportaba como un monómero en solución. Estos resultados indicaron que las interacciones homofílicas a través de los cuatro dominios N-terminales eran débiles y que posiblemente se requiere el dominio 5 para obtener interacciones de adhesión relevantes.

Tras varios pasos cromatográficos se obtuvo una proteína de una gran pureza (Fig.4.13F), con la que pudimos llevar a cabo ensayos de cristalización. La proteína era altamente estable y podía ser conservada a 4°C durante varias semanas o congelada a -20°C durante largos periodos de tiempo sin sufrir degradación.

### 5.3.2. Variabilidad en la cristalización de IC5-4D y resolución de la estructura

La variabilidad cristalográfica observada con IC5-4D en condiciones muy similares que indican que esta proteína puede mediar diversas interacciones intermoleculares, algunas dependientes del pH del medio y consecuentemente de la carga de la proteína. El elevado número de formas cristalinas que se obtienen con esta proteína no se ha observado con otros ICAMs, por lo que pueden estar relacionadas con la peculiar distribución de cargas en los dominios de ICAM-5. De hecho en las formas X2 y X3 se aprecian contactos cristalográficos entre los dominios N-terminales básicos y los C-terminales ácidos (Fig. 4.16).

Se decidió iniciar el proceso de resolución de la estructura a partir de la forma X3 que difractaba a más alta resolución, a pesar de que la obtención de los cristales fue más tardía y mayor dificultad que para las otras formas. No obstante, a partir de los gráficos de Wilson se observó que los factores de temperatura de la forma X3 eran relativamente altos ( $60 \text{ \AA}^2$ ), que correlaciona con el alto contenido de solvente de los cristales (60-70%). Para la determinación de la estructura se siguió inicialmente el método de reemplazamiento isomorfo múltiple (MIR) mediante la búsqueda al azar de derivado de átomos pesados, principalmente porque IC5-4D sólo cuenta con un residuo de metionina que pueda ser sustituido por seleno-metionina. La búsqueda de derivados fue extremadamente tediosa, posiblemente porque los altos factores de temperatura dificultaron la identificación de sitios de unión de átomos pesados a la proteína. En ningún caso se obtuvieron sitios de unión evidentes, que resultan en señales altas en los mapas de Patterson. Inicialmente se utilizaron compuestos con cargas positivas (Sm y Yb) y negativas (Br y  $\text{Pt}(\text{SCN})_6$ ) que se unen electrostáticamente a las proteínas, ya que se esperaba que estos átomos podrían unirse a los dominios cargados de ICAM-5 a pH neutro. Una vez obtenidos sitios de unión iniciales se colectaron datos cristalográficos de derivados seleccionados al azar, que permitieron obtener un derivado que se coordina más fuertemente a las proteínas ( $\text{Pt}(\text{CN})_4$ ), que generó mapas de mayor calidad, en los que se podían identificar las moléculas de IC5-4D.

Alternativamente, una vez publicada la estructura de los dominios 3, 4 y 5 de ICAM-1 (Yang Y 2004) decidimos abordar la determinación de la estructura por el método de reemplazo molecular. La determinación de la estructura con la consiguiente localización de la molécula en la unidad asimétrica requirió la búsqueda de soluciones para los dominios individuales. La solución seleccionada presentó pocas dudas sobre su

veracidad, ya que los dominios aparecían concatenados, con los extremos C y N terminales a la distancia esperada. Asimismo, los mapas de densidad mostraron densidad electrónica para los carbohidratos ausentes en el modelo. Finalmente se combinaron las fases que provenían de las soluciones de reemplazo molecular con las obtenidas a partir de los derivados de átomos pesados con objeto de reducir los errores presentes en el modelo. La combinación de fases mejoró los mapas de densidad electrónica, particularmente en la conexión entre los dominios de inmunoglobulina.

Seguidamente se procedió al refinamiento de la estructura inicialmente mediante reconstrucción manual del modelo y seguidamente por minimización de energía con el programa CNS (Brünger AT 1998). El modelo se sometió a varios procesos de refinamiento manual y automático, que permitió refinar parcialmente el modelo. No obstante, el proceso de refinamiento ha sido complicado, posiblemente por el alto factor de temperatura de la única molécula presente en la unidad asimétrica de los cristales, relacionado con el alto contenido de solvente y los escasos contactos cristalinos observados en la forma X3. En la actualidad el modelo se ha refinado hasta valores de  $R_{\text{libre}}$  de aproximadamente 30% usando datos a 3Å de resolución (no mostrado en esta Tesis).

Las formas cristalográficas X1 y X2 se resolvieron por el método de reemplazo molecular usando dos modelos con los dominios D1-D2 y D3-D4. Se continúa trabajando en el refinamiento de estas estructuras en la actualidad.

### **5.3.3. Estructura de los cuatro dominios N-terminales de ICAM-5**

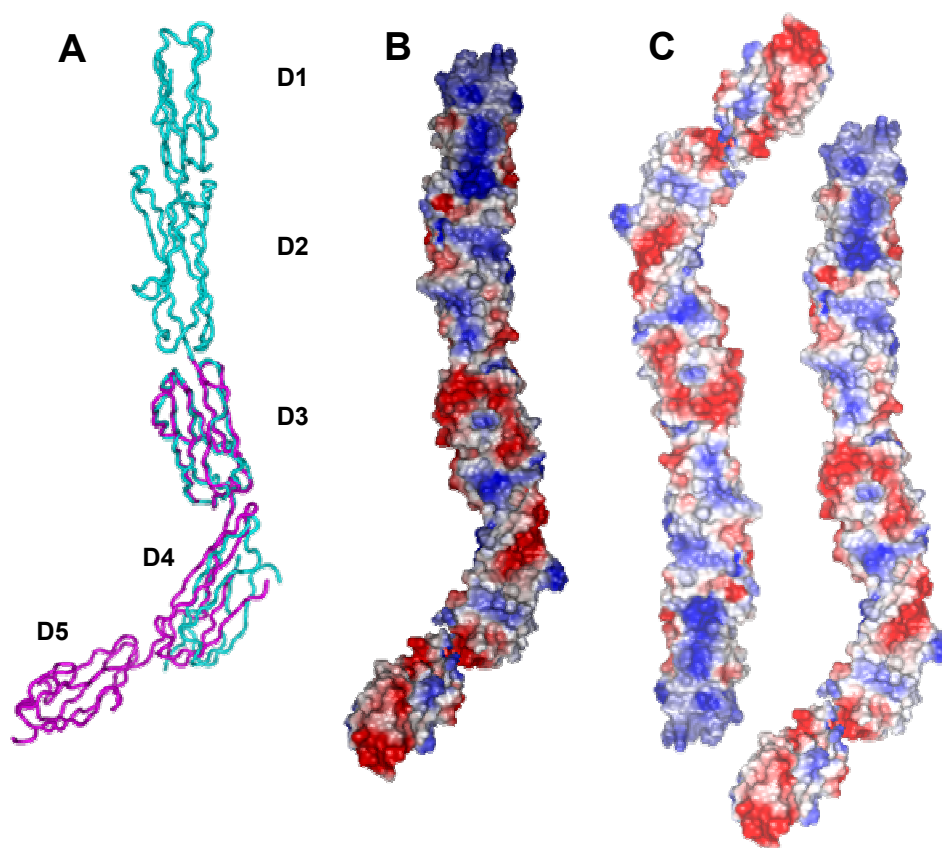
La estructura de los dominios de ICAM-5 presentados presenta gran similitud con los dominios homólogos de ICAM-1. Las mayores diferencias aparecen en bucles del primer lazo DE implicado en dimerización de ICAM-1, y segundo dominio, lazo C'E que es más extenso en ICAM-5 y que se había observado que presentaba alta flexibilidad en la estructura de ICAM-2. Las N-glicosilaciones aparecen conservadas principalmente en el dominio 2 de ICAM-1 e ICAM-5. Este receptor contienen la glicosilación conservada en la hebra B del dominio 1, que interacciona con la cadena lateral de un triptófano, tal y como se observó en ICAM-2. El dominio 4, que aparece incompleto en la figura 4.19, se ha modelado reconstruido completamente durante los últimos pasos de refinamiento. Este modelo es del mismo tipo que el dominio 2 (grupo

I2) y ambos presentan unos bucles FG particularmente extensos que los conectan con el dominio precedente. Por otra parte, los dominios 1 y 3 son del tipo I1, lo que sugiere que el módulo de cuatro dominios ha aparecido por la duplicación de un tandem de dominios I1-I2.

La resolución de varias estructuras ha permitido identificar cierta flexibilidad entre dominios. Parece que el grado de flexibilidad es mayor entre los tres primeros dominios que en el tandem D3-D4, que parece relativamente fijo. Asimismo, mientras la disposición de los dominios D1 a D3 es relativamente recta, D4 se dispone relativamente oblicuo a D3. La variabilidad estructural de IC5-4D parece dependiente del pH ya que la forma X1 cristalizada a bajo pH presenta mayor flexibilidad que las formas cristalizadas a alto pH (Fig. 4.17A).

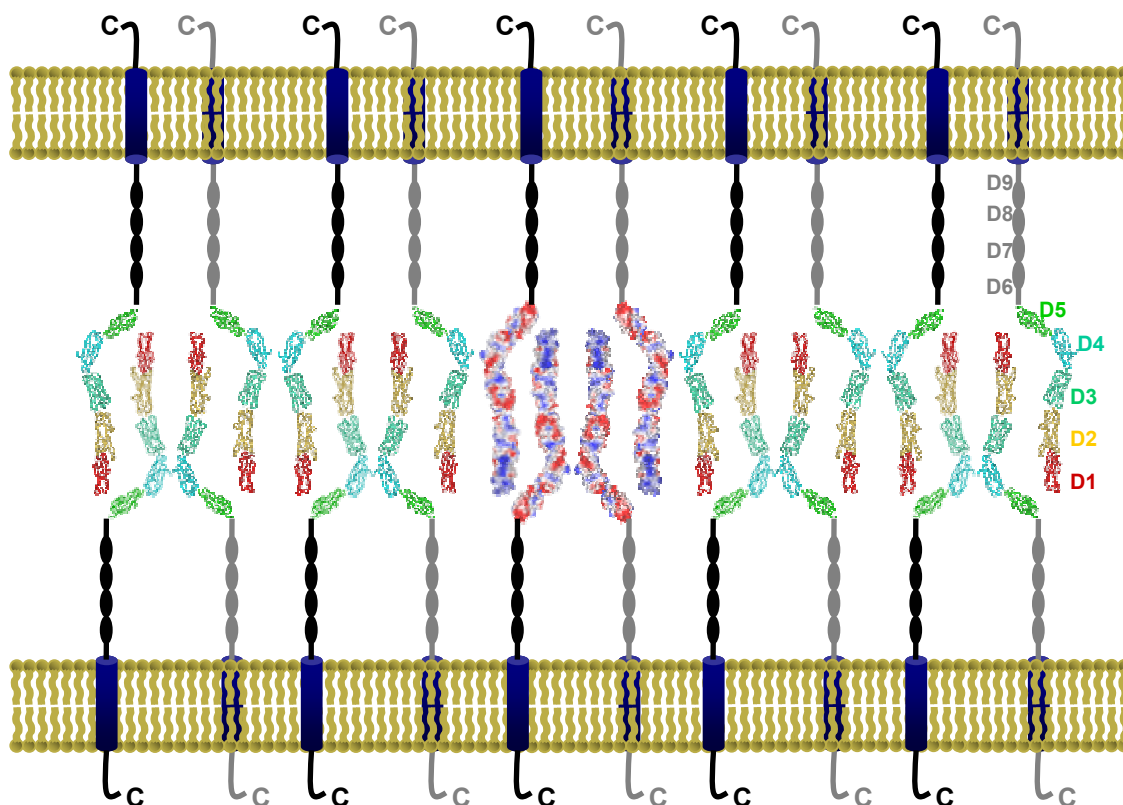
#### **5.3.4. Modelo para la interacción homofílica en ICAM-5**

El objetivo último del análisis estructural de ICAM-5 es entender los mecanismos de adhesión homofílica específicos de este receptor celular, del que este trabajo de Tesis es un primer paso para la consecución de este objetivo (Tian L 2000). Las interacciones de adhesión homofílica de ICAM-5 implican a los cinco dominios N-terminales del receptor (Tian L 2000). El dominio 5 es fuertemente ácido, al igual que los dominios 3 y 4, los cuales deben de interaccionar electrostáticamente con los dominios 1 y 2, fuertemente básicos. Si bien en la actualidad no tenemos la estructura de la proteína con los dominios 1 a 5, podemos utilizar la estructura de los dominios 3 a 5 de ICAM-1 para modelar IC5-5D (Fig. 5.1.A). La superposición del dominio 3 de ICAM-1 e ICAM-5 muestra que el módulo D4-D5 se dispone relativamente oblicuo al módulo D1-D3 (Fig. 5.1). Esta superposición muestra que el ángulo de aproximadamente 50° entre D3 y D4 es muy similar en ICAM-1 e ICAM-5. La molécula de ICAM-5 tiene aspecto de palo de jockey, con la empuñadura (D1) básica y la cabeza o zona distal ácida (Fig. 5.1.B ). La distribución de cargas sugiere una posible interacción intermolecular cabeza-empuñadura entre dos moléculas dispuestas en orientaciones opuestas (Fig. 5.1.C).



**Fig.5.1. Modelo de ICAM-5 D1-D5.** A. Superposición de las estructuras tridimensionales de ICAM-5 D1-D4 (azul) con ICAM-1 D3-D5 (morado), alineadas por el dominio D3. B. Esquema de distribución de cargas de la molécula ICAM-5 D1-D5, los dominios 1 y 2 son fuertemente básicos, mientras que los dominios 3, 4 y 5 son fundamentalmente ácidos. C. Modelo de interacción intermolecular entre dos moléculas de ICAM-5 dispuestas en orientación opuesta. (Las imágenes se prepararon usando el programa PyMol, DeLano Scientific LLC).

En la estructura de IC5-4D en su forma cristalina X2 (Fig.4.16), se observa una distribución de las moléculas similar a la mostrada en la Fig 5.1.C. Sin embargo, en esta forma cristalina aparecen dos moléculas en la unidad asimétrica, que forman un dímero en el que ambas moléculas se disponen en la misma orientación. Esta disposición es representativa de un dímero del receptor sobre la misma superficie celular. Si bien no hemos observado dimerización de IC5-4D en solución, si tenemos datos preliminares sobre la formación de dímeros por IC5-5D, que sugieren dimerización de ICAM-5 sobre la superficie celular, tal y como se ha descrito para ICAM-1.



**Fig.5.2. Modelo teórico de interacción homofílica entre moléculas de ICAM-5 situadas en células opuestas.** Las moléculas de ICAM-5 se sitúan sobre la superficie celular formando dímeros en *cis* los cuales interaccionan en *trans* con otros dímeros situados en células opuestas, debido a las diferencias de carga existentes entre los dominios D1,D2 (básicos) y D3,D4, D5 (ácidos).

A partir de estos datos hemos elaborado un modelo para la interacción homofílica mediada por ICAM-5 (Fig. 5.2). Según este modelo, dos dímeros de ICAM-5 sobre células opuestas interaccionan entre sí tal y como se describe en la figura 5.1. Asimismo, la proteína puede formar una estructura tipo cremallera en las uniones intercelulares. Este modelo va a ser validado por mutagénesis y mediante la determinación de la estructura de los 5 dominios N-terminales del receptor.







## **CAPITULO 6: CONCLUSIONES**



## 6. CONCLUSIONES

- Los azúcares de tipo N participan en la estructura y función de la subfamilia de receptores ICAMs, interaccionando con residuos proteicos, modulando la accesibilidad de las superficies proteicas o participando en el reconocimiento de ligandos.
- Los carbohidratos facilitan el plegamiento del dominio N-terminal de ICAM-1 aislado, dependiendo de la localización de los mismos.
- Los N-glicanos en la subfamilia de los ICAMs proporcionan la especificidad de unión a la lectina DC-SIGN.
- La actividad de unión a DC-SIGN está relacionada con la presencia de carbohidratos con alto contenido en manosa, unidos específicamente al segundo dominio N-terminal de ICAM-2 e ICAM-3.
- No hay interacciones significativas entre residuos proteicos de ICAMs y DC-SIGN.
- Los cuatro dominios N-terminales de ICAM-5 presentan cierta variabilidad cristalográfica, que puede reflejar variabilidad en la conformación de la proteína en solución.
- La estructura de los cuatro dominios N-terminales de ICAM-5 es similar a la de los dominios homólogos de ICAM-1.
- Existe una mayor flexibilidad entre los tres dominios N-terminales que entre los dominios 3 y 4, que presentan un módulo relativamente rígido y doblado.
- Los cuatro dominios N-terminales de ICAM-5 adoptan una disposición de “palo de jockey”, con la empuñadura básica y la cabeza ácida.
- La conformación modelada para una molécula con los 5 dominios N-terminales de ICAM-5 sugiere una interacción de adhesión homofílica de naturaleza electrostática.



## **CAPITULO 7: BIBLIOGRAFÍA**





## 7. BIBLIOGRAFIA

- Alberts B, B. D., Lewis J, Raff M, Roberts K, Watson JD (1994). MOLECULAR BIOLOGY OF THE CELL. New York, Garland Publishing Inc.
- Ashford DA, A. C., Gamble VM, Mackay DJ, Rademacher TW, Williams PJ, Dwek RA, Barclay AN, Davis SJ, Somoza C, Ward HA and Williams AF (1993). "Site-specific glycosylation of recombinant rat and human soluble CD4 variants expressed in Chinese hamster ovary cells." Journal of Biological Chemistry **268**(5): 3260-7.
- Bella J, K. P., Marlor CW, Greve JM, Rossmann MG. (1998). "The structure of the two amino-terminal domains of human ICAM-1 suggests how it functions as a rhinovirus receptor and as an LFA-1 integrin ligand." Proceedings of the National Academy of Sciences **95**(8): 4140-5.
- Bella J, R. M. (1999). "Review: Rhinoviruses and Their ICAM Receptors." Journal of Structural Biology **128**: 69-74.
- Berg JM, T. J., Stryer L. (2002). BIOCHEMISTRY. New York, W.H. Freeman and Company.
- Bergfors, T. M. (1999). Protein crystallization. Techniques, strategies, and tips, International University Line.
- Bleijds DA, G. T., Figdor CG, van Kooyk Y. (2001). "DC-SIGN and LFA-1: a battle for ligand." Trends in Immunology **22**: 457-63.
- Bretscher, P. (1999). "A two-step, two-signal model for the primary activation of precursor helper T cells." Proceedings of the National Academy of Sciences **96**: 185-90.
- Brünger AT, P. D. A., G. Marius Clore, Warren L. DeLano, Piet Gros, Ralf W. Grosse-Kunstleve, Jian-Sheng Jiang, John Kuszewski, Michael Nilges, Navraj S. Pannu, Randy J. Read, Luke M. Rice, Thomas Simonson, and Gregory L. Warren. (1998). "Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination." Acta Cryst. D **54**: 905-921.
- Casasnovas JM, B. J., Springer TA. (1998). "The domain structure of ICAM-1 and the kinetics of binding to rhinovirus." Journal of Virology **72**: 6244-46.
- Casasnovas JM, P. C., Springer TA. (1999). "Lymphocyte function-associated antigen-1 binding residues in intercellular adhesion molecule-2 (ICAM-2) and the integrin binding surface in the ICAM subfamily." Proceedings of the National Academy of Sciences **96**: 3017-3022.
- Casasnovas JM, S. T., Liu JH, Harrison SC, Wang JH. (1997). "Crystal structure of ICAM-2 reveals a distinctive integrin recognition surface." Nature **387**: 312-5.
- Casasnovas JM, S. T., Liu JH, Wang JH, Springer TA. (1998). "A dimeric crystal structure for the N-terminal two domains of intercellular adhesion molecule-1." Proceedings of the National Academy of Sciences **95**: 4134-9.
- Collaborative Computational Project, N. (1994). "The CCP4 Suite: Programs for Protein Crystallography." Acta Cryst. D **50**: 760-763.
- Chang KC, T. D., Hammer DA. (2000). "The state diagram for cell adhesion under flow: leukocyte rolling and firm adhesion." Proceedings of the National Academy of Sciences **97**(21): 11262-7.
- Drickamer, K. (1999). "C-type lectin-like domains." Current Opinion in Structural Biology **9**(5): 585-90.

- DuBridge RB, T. P., Hsia HC, Leong PM, Miller JH, Calos MP. (1987). "Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system." Molecular Cell Biology **7**: 379-87.
- Dustin ML, S. T. (1989). "T-cell receptor cross-linking transiently stimulates adhesiveness through LFA-1." Nature **341**: 619-24.
- Dwek, R. (1995). "Glycobiology: more functions for oligosaccharides." Science **269**(5228): 1234-5.
- Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallogr D Biol Crystallogr. **62**: 72-82.
- Feinberg H, M. D., Drickamer K, Weis WI. (2001). "Structural basis for selective recognition of oligosaccharides by DC-SIGN and DC-SIGNR." Science **294**: 2163-6.
- Figdor CG, v. K. Y., Adema GJ. (2002). "C-type lectin receptors on dendritic cells and Langerhans cells." Nat Rev Immunol. **2**: 77-84.
- Funatsu O, S. T., Kotovuori P, Gahmberg CG, Ikekita M, Furukawa K. (2001). "Structural study of N-linked oligosaccharides of human intercellular adhesion molecule-3 (CD50)." European Journal of Biochemistry **268**: 1020-9.
- Garboczi DN, U. U., Ghosh P, Seth A, Kim J, VanTienhoven EA, Biddison WE, Wiley DC (1996). "Assembly, specific binding, and crystallization of a human TCR-alpha-beta with an antigenic Tax peptide from human T lymphotropic virus type 1 and the class I MHC molecule HLA-A2." J Immunol **157**: 5403-10.
- Geijtenbeek TB, K. D., Bleijs DA, van Vliet SJ, van Duijnhoven GC, Grabovsky V, Alon R, Figdor CG, van Kooyk Y. (2000). "DC-SIGN-ICAM-2 interaction mediates dendritic cell trafficking." Nature Immunology **1**: 353-7.
- Geijtenbeek TB, T. R., van Vliet SJ, van Duijnhoven GC, Adema GJ, van Kooyk Y, Figdor CG. (2000). "Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses." Cell **100**: 491-4.
- Grakoui A, B. S., Sumen C, Davis MM, Shaw AS, Allen PM, Dustin ML. (1999). "The immunological synapse: a molecular machine controlling T cell activation." Science **285**(5425): 221-7.
- Greve JM, D. G., Meyer AM, Forte CP, Yost SC, Marlor CW, Kamarck ME, McClelland A (1989). "The major human rhinovirus receptor is ICAM-1." Cell **56**: 839-47.
- Guillemot, F. (2005). "Cellular and molecular control of neurogenesis in the mammalian telencephalon." Current Opinion in Cell Biology **17**(6): 639-647.
- Guo Y, F. H., Conroy E, Mitchell DA, Alvarez R, Blixt O, Taylor ME, Weis WI, Drickamer K. (2004). "Structural basis for distinct ligand-binding and targeting properties of the receptors DC-SIGN and DC-SIGNR." Nat Struct Mol Biol. **11**: 591-8.
- Harpaz Y, C. C. (1994). "Many of the immunoglobulin superfamily domains in cell adhesion molecules and surface receptors belong to a new structural set which is close to that containing variable domains." J Mol Biol. **238**: 528-30.
- Hynes, R. (1992). "Integrins: versatility, modulation, and signaling in cell adhesion." Cell **69**: 11-25.
- Hynes, R. (1999). "Cell adhesion: old and new questions." Trends in Cell Biology **9**(12): M33-37.
- Jiménez D, R.-N. P., Springer TA, Casasnovas JM (2005). "Contribution of N-linked glycans to the conformation and function of intercellular adhesion molecules (ICAMs)." The Journal of Biological Chemistry **280**: 5854-5861.

- Jones TA, Z. J., Cowan SW, Kjeldgaard. (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Cryst. A **47**: 110-119.
- Kabsch, W. (1993). "Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants." J. Appl. Cryst. **26**: 795-800.
- Klickstein LB, Y. M., Fougerolles AR, Springer TA. (1996). "Localization of the binding site on intercellular adhesion molecule-3 (ICAM-3) for lymphocyte function-associated antigen 1 (LFA-1)." Journal of Biological Chemistry **271**(39): 23920-7.
- Kolatkar PR, B. J., Olson NH, Bator CM, Baker TS, Rossmann MG (1999). "Structural studies of two rhinovirus serotypes complexed with fragments of their cellular receptor." EMBO J **18**: 6249-59.
- Kwong PD, D. M., Casper DJ, Cicala C, Leavitt SA, Majeed S, Steenbeke TD, Venturi M, Chaiken I, Fung M, Katinger H, Parren PW, Robinson J, Van Ryk D, Wang L, Burton DR, Freire E, Wyatt R, Sodroski J, Hendrickson WA, Arthos J. (2002). "HIV-1 evades antibody-mediated neutralization through conformational masking of receptor-binding sites." Nature **420**(6916): 678-82.
- Larson RS, C. A., Berman L, Springer T (1989). "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily." J Cell Biol. **108**: 703-712.
- Lee JO, B. L., Arnaout MA, Liddington RC (1995). "Two conformations of the integrin A-domain (I-domain): a pathway for activation?" Structure. **3**: 1333-40.
- Lee JO, R. P., Arnaout MA, Liddington R (1995). "Crystal structure of the A domain from the alpha subunit of integrin CR3 (CD11b/CD18)." Cell **80**: 6318.
- Leslie, A. (1992). "Jnt CCP4/ESF-EACMB." Newslett. Protein Crystallogr. **26**.
- Luster AD, A. R., von Adrian UH (2005). "Immune cell migration in inflammation: present and future therapeutic targets." Nature Immunology **6**: 1182-90.
- Martin S, C. J., Staunton DE, Springer TA (1993). "Efficient neutralization and disruption of rhinovirus by chimeric ICAM-1/immunoglobulin molecules." J Virol. **67**: 3561-8.
- Mizuno T, Y. Y., Inazawa J, Kagamiyama H, Mori K (1997). "cDNA cloning and chromosomal localization of the human telencephalin and its distinctive interaction with lymphocyte function-associated antigen-1." The Journal of Biological Chemistry **272**: 1156-1163.
- Mizushima S, N. S. (1990). "pEF-BOS, a powerful mammalian expression vector." Nucleic Acids Res. **18**: 5322.
- Mori K, F. S., Watanabe Y, Obata K, Hayaishi O. (1987). "Telencephalon-specific antigen identified by monoclonal antibody." Proceedings of the National Academy of Sciences **84**(11): 3921-5.
- Paniagua R, N. M., Sesma P, Álvarez-Uría M, Fraile B, Anadón R, Sáez FJ. (2003). BIOLOGÍA CELULAR. Madrid, McGRAW-HILL-INTERAMERICANA DE ESPAÑA.
- Patarroyo M, P. J., Rincon J, Timonen T, Lundberg C, Lindbom L, Asjo B, Gahmberg CG (1990). "Leukocyte-cell adhesion: a molecular process fundamental in leukocyte physiology." Immunol Rev. **114**(67-108).
- Paulson, J. (1989). "Glycoproteins: what are the sugar chains for?" Trends in biochemical sciences. **14**(7): 272-6.
- Pear WS, N. G., Scott ML, Baltimore D. (1993). "Production of high-titer helper-free retroviruses by transient transfection." Proceedings of the National Academy of Sciences **90**: 8392-6.

- Perez-Moreno M, J. C., Fuchs E (2003). "Sticky business: orchestrating cellular signals at adherens junctions." *Cell* **21**: 535-48.
- Qu A, L. D. (1995). "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin." *Proceedings of the National Academy of Sciences* **92**: 10277-81.
- Relloso M, P.-K. A., Pello OM, Rodriguez-Fernandez JL, de la Rosa G, Longo N, Navarro J, Munoz-Fernandez MA, Sanchez-Mateos P, Corbi AL. (2002). "DC-SIGN (CD209) expression is IL-4 dependent and is negatively regulated by IFN, TGF-beta, and anti-inflammatory agents." *Journal of Immunology* **168**: 2634-43.
- Rodgers, D. (1994). "Cryocrystallography." *Structure*. **15**: 1135-40.
- Roitt I, B. J., Male D (1998). *IMMUNOLOGY*. London, Mosby International Ltd.
- Rudd PM, E. T., Cresswell P, Wilson IA, Dwek RA. (2001). "Glycosylation and the immune system." *Science* **291**(5512): 2370-6.
- Shimaoka M, X. T., Liu JH, Yang Y, Dong Y, Jun CD, McCormack A, Zhang R, Joachimiak A, Takagi J, Wang JH, Springer TA. (2003). "Structures of the alpha L I domain and its complex with ICAM-1 reveal a shape-shifting pathway for integrin regulation." *Cell* **112**(1): 99-111.
- Song G, Y. Y., Liu JH, Casasnovas JM, Shimaoka M, Springer TA, Wang JH (2005). "An atomic resolution view of ICAM recognition in a complex between the binding domains of ICAM-3 and integrin alphaLbeta2." *Proceedings of the National Academy of Sciences* **102**: 3366-71.
- Springer, T. (1990). "Adhesion receptors of the immune system." *Nature* **346**: 425-433.
- Springer, T. (1994). "Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm." *Cell* **76**(2): 301-14.
- Stanley, P. (1989). "Chinese hamster ovary cell mutants with multiple glycosilation defects for production of glycoproteins with minimal carbohydrate heterogeneity." *Molecular and Cellular Biology* **9**: 377-383.
- Staunton DE, D. M., Erickson HP, Springer TA. (1990). "The arrangement of the immunoglobulin-like domains of ICAM-1 and the binding sites for LFA-1 and rhinovirus." *Cell* **61**(2): 243-54.
- Staunton DE, M. V., Rothlein R, Barton R, Marlin SD, Springer TA (1989). "A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses." *Cell* **56**: 849-53.
- Tian L, K. P., Yoshihara Y, Mori K, Gallatin WM, Carpen O, Gahmberg CG. (2000). "Binding of T lymphocytes to hippocampal neurons through ICAM-5 (telencephalin) and characterization of its interaction with the leukocyte integrin CD11a/CD18." *European Journal of Immunology* **30**: 810-818.
- Tian L, N. H., Kilgannon P, Yoshihara Y, Mori K, Andersson LC, Kaukinen S, Rauvala H, Gallatin WM, Gahmberg CG. (2000). "Intercellular Adhesion Molecule-5 Induces Dendritic Outgrowth by Homophilic Adhesion." *The Journal of Cell Biology* **150**: 243-252.
- Vestweber D, B. J. (1999). "Mechanisms that regulate the function of the selectins and their ligands." *Physiological Reviews* **79**(1): 181-213.
- Wyss DF, C. J., Li J, Knoppers MH, Willis KJ, Arulanandam AR, Smolyar A, Reinherz EL, Wagner G. (1995). "Conformation and function of the N-linked glycan in the adhesion domain of human CD2." *Science* **269**(5228): 1273-8.
- Yang Y, J. C., Liu JH, Zhang R, Joachimiak A, Springer TA, Wang JH. (2004). "Structural basis for dimerization of ICAM-1 on the cell surface." *Molecular Cellular* **14**: 269-276.





## **ANEXO 1**



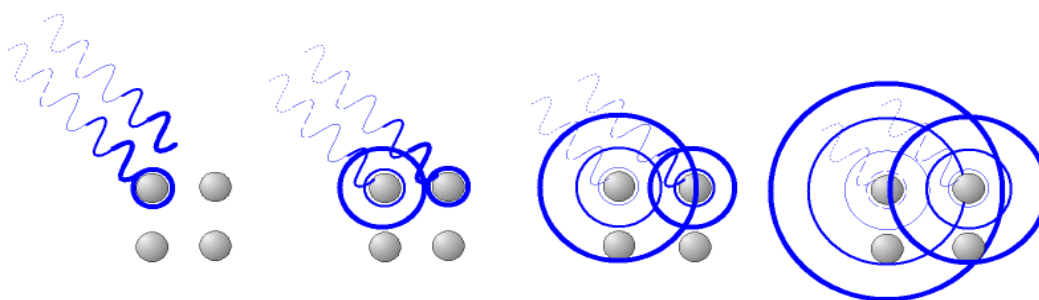


## ANEXO 1: PRINCIPIOS DE LA DETERMINACIÓN DE ESTRUCTURAS PROTEICAS MEDIANTE CRISTALOGRAFÍA DE RAYOS X

### A1.1- Introducción

El método consiste en obtener un mapa de densidad electrónica tridimensional de una proteína a partir de los datos de difracción de rayos X.

Los rayos X, que son ondas electromagnéticas, interactúan con los electrones de las moléculas que forman el cristal. Los componentes eléctricos y magnéticos de la onda incidente ejercen una fuerza sobre los electrones, que los hace oscilar con la misma frecuencia que ella. Los electrones oscilantes emiten una radiación de la misma frecuencia (longitud de onda) que la radiación incidente en todas las direcciones. Este fenómeno se conoce como dispersión coherente, y es la única que se utiliza en cristalografía (Fig.A.1.1).



**Fig.A1.1.** Cuando los rayos X alcanzan un átomo interactúan con sus electrones exteriores. Estos remiten la radiación electromagnética incidente en diferentes direcciones y con pequeños cambios en su frecuencia. Este fenómeno se conoce como dispersión de Rayleigh. Los rayos X remitidos desde átomos cercanos interfieren entre sí constructiva o destructivamente. Este es el fenómeno de la difracción. En el presente diagrama se esquematizan los rayos X que inciden sobre un cristal. Los átomos superiores remiten la radiación tras ser alcanzados por ella. Los puntos en los que la radiación se superpone constructivamente se muestran como la zona de intersección de los anillos. Se puede apreciar que existen ángulos privilegiados en los cuales la interferencia es constructiva, en este caso a la derecha, con un ángulo en torno a 45°.

La teoría de dispersión de rayos X a través de un cristal se basa en una aproximación según la cual los electrones de los átomos del cristal se comportan como electrones libres. Cuando un haz de rayos X de dirección  $s_0$  golpea un cristal, la dispersión tiene lugar en cada electrón del cristal. La onda dispersada resultante es la suma de todas las ondas, cada una correspondiente a un electrón en una posición particular del cristal. Los electrones en los átomos forman una “nube” con una densidad de carga  $\rho(\mathbf{r})$  en cada

posición  $\mathbf{r}(x, y, z)$  con respecto al origen del cristal. En cada elemento de volumen  $\Delta\mathbf{r}$ , la carga neta contenida es igual a  $\rho(\mathbf{r})\Delta\mathbf{r}$ . La expresión para definir la onda total dispersada por un cristal ( $\mathbf{F}$ ) se obtiene sumando la dispersión de la nube electrónica para cada elemento de volumen. De esta forma, se puede expresar la densidad electrónica como una función continua. Así, el sumatorio sobre el volumen del cristal se reemplaza por una integral para expresar la dispersión total:

$$\mathbf{F}(\mathbf{S}) = \int \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \bullet \mathbf{S}) d\mathbf{r}$$

La variable  $\mathbf{S}$  se llama vector de dispersión, y su amplitud,  $|\mathbf{S}|$ , es una función del ángulo de dispersión, que se define por las direcciones de la onda incidente ( $\mathbf{s}_0$ ) y dispersada ( $\mathbf{s}$ ). El valor de  $|\mathbf{S}|$  varía entre 0 y  $2/\lambda$ . El vector  $\mathbf{S}$  se puede describir en un sistema de coordenadas llamado “espacio recíproco”.

La onda total dispersada por un cristal ( $\mathbf{F}$ ) se denomina factor de estructura, porque su magnitud y su fase son determinadas por la estructura molecular y el ordenamiento de las moléculas de proteína en el cristal. La ecuación que relaciona el factor de estructura con la densidad electrónica es equivalente a una transformada de Fourier. Así, se puede decir que el factor de estructura es la transformada de Fourier de la densidad electrónica del cristal. De este modo, si la distribución de la densidad electrónica en el cristal,  $\rho(\mathbf{r})$ , es conocida, entonces se pueden calcular los factores de estructura para cada ángulo de dispersión permitido, y de este modo generar el patrón de difracción completo del cristal. El problema en cristalografía es que es la distribución de la densidad electrónica lo que no conocemos, y los factores de estructura los podemos calcular a partir de los datos de difracción medidos, luego tenemos que calcular la transformada de Fourier inversa:

$$\rho(\mathbf{r}) = 1/V \sum_s \mathbf{F}(\mathbf{S}) \exp(-2\pi i \mathbf{r} \bullet \mathbf{S})$$

La integral se reemplaza por el sumatorio sobre el espacio recíproco ya que la difracción ocurre solo en determinadas direcciones. Solo para ciertos valores del vector de dispersión  $\mathbf{S}$ , la dispersión para cada elemento de volumen del cristal se puede combinar de forma constructiva para originar un factor de estructura con una magnitud lo suficientemente grande para ser medida.

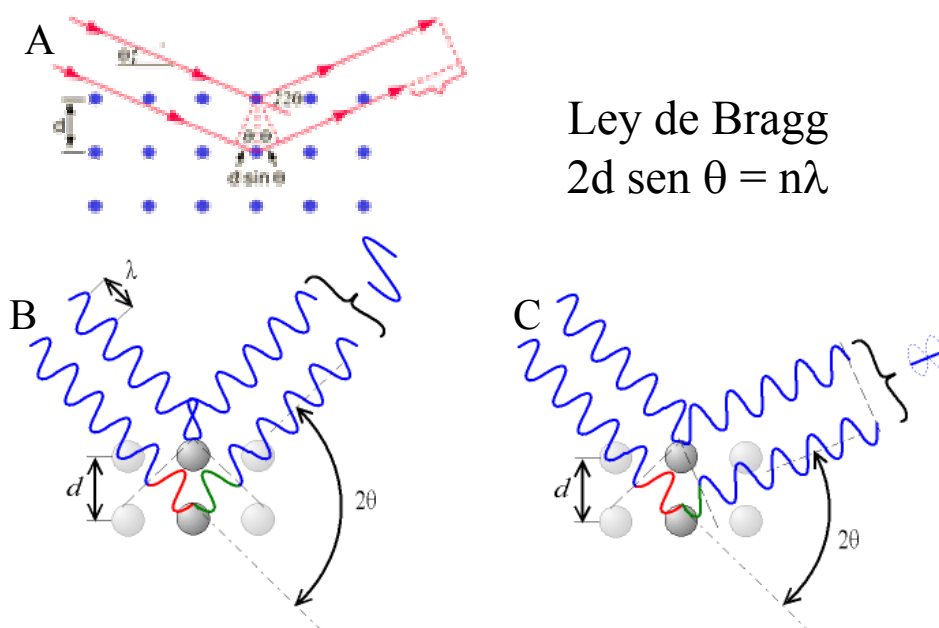
Los valores de  $\mathbf{S}$  permitidos son aquellos que satisfacen las siguientes condiciones:

$$\mathbf{a} \cdot \mathbf{S} = h \quad \mathbf{b} \cdot \mathbf{S} = k \quad \mathbf{c} \cdot \mathbf{S} = l$$

donde **a**, **b** y **c** son los ejes de la celda unidad, y *h*, *k* y *l* son números enteros. La constante *V* es el volumen de la celda unidad. La ecuación de la densidad electrónica implica que si se miden todos los factores de estructura del espacio recíproco se puede calcular la distribución de la densidad electrónica del cristal.

### A1.2.- La ley de Bragg

La ley de Bragg es la forma más extendida de explicar el fenómeno de la difracción de cristales por medio de rayos X. Según Bragg, un rayo X incidente sobre una red de planos con un ángulo de incidencia  $\theta$  es reflejado por el plano con el mismo ángulo. Esto corresponde a un ángulo de dispersión igual a  $2\theta$  (Fig.A1.2).



**Fig.A1.2. Ley de Bragg.** Un haz de rayos X incidente sobre una red de planos con un ángulo de incidencia  $\theta$  es reflejado por el plano con el mismo ángulo (A). Cuando el ángulo de desviación es  $2\theta$ , el cambio de fases de las ondas produce interferencia constructiva (B) o destructiva (C).

La condición de Bragg para observar una onda “reflejada” es que la diferencia entre ondas “reflejadas” por planos adyacentes, en un conjunto de redes de planos, sea un número entero de longitudes de onda. Esto se expresa en una ecuación conocida como ley de Bragg (Fig.A1.2). Donde *n* es cualquier número entero y *d* es la distancia

interplanar en un conjunto de redes de planos  $h,k,l$ . Cuando la diferencia entre planos es un número entero de longitudes de onda la diferencia de fase es periódica a lo largo de  $360^\circ$ . De este modo los puntos en planos sucesivos dispersan electrones en fase y las ondas se combinan de forma constructiva. Cada vector de dispersión  $\mathbf{S}$  es perpendicular a un conjunto de redes de planos  $(hkl)$ , la magnitud de  $\mathbf{S}$  es igual a la inversa de la distancia interplanar.

### **A.1.3.- Red recíproca**

Los puntos finales de todos los vectores de dispersión  $\mathbf{S}$   $(hkl)$  forman una red de vectores de traslación  $\mathbf{a}^*$ ,  $\mathbf{b}^*$ ,  $\mathbf{c}^*$ , que está relacionada con los ejes  $\mathbf{a}, \mathbf{b}, \mathbf{c}$  de la red del cristal. Los ejes  $\mathbf{a}^*$ ,  $\mathbf{b}^*$  y  $\mathbf{c}^*$  tienen dimensiones recíprocas a  $\mathbf{a}$ ,  $\mathbf{b}$  y  $\mathbf{c}$ . Por esto, la red que forman se denomina red recíproca. Cada reflexión en un conjunto de planos cristalinos  $hkl$  corresponde a un punto con coordenadas  $h, k, l$  en el espacio recíproco, y el vector de dispersión se puede expresar como:

$$\mathbf{S}(hkl) = h \mathbf{a}^* + k \mathbf{b}^* + l \mathbf{c}^*$$

La simetría de la red recíproca está relacionada con la simetría de la red del cristal. El conjunto de puntos de difracción registrado por el detector de rayos x es una imagen de la red recíproca. Por lo tanto, analizando el patrón de difracción se puede observar la simetría del cristal y calcular las dimensiones de la celda.

### **A1.4.- Resumen**

El resultado de la interacción de un haz de rayos X con un cristal es la dispersión de la radiación en ciertas direcciones concretas del espacio. Si se coloca un detector perpendicular al rayo incidente, los rayos X difractados producirán en él un conjunto de puntos, cada uno de ellos correspondiente a un haz de rayos X difractado por el cristal y de una intensidad determinada. Esto es lo que se considera un patrón de difracción. Hay que remarcar que los puntos de difracción no son posiciones de los átomos individuales en el cristal, pero cada punto de difracción lleva información sobre los electrones (y átomos) en el cristal.

La expresión del vector de dispersión **S** en términos de la red recíproca, y del vector de posición **r** en términos de coordenadas cartesianas x, y, z nos da:

$$\mathbf{F}(\mathbf{hkl}) = V \int \rho(\mathbf{xyz}) \exp \{ 2\pi i (\mathbf{xh} + \mathbf{yk} + \mathbf{zl}) \} d\mathbf{x} d\mathbf{y} d\mathbf{z}$$

y

$$\rho(\mathbf{xyz}) = 1/V \sum_{\mathbf{hkl}} \mathbf{F}(\mathbf{hkl}) \exp \{ -2\pi i (\mathbf{xh} + \mathbf{yk} + \mathbf{zl}) \}$$

para las ecuaciones de factores de estructura y densidad electrónica, respectivamente.

Desgraciadamente, en estudios de cristalografía de rayos X, no hay forma de medir **F(S)** directamente. **F** es un número complejo con una amplitud y una fase. Experimentalmente, todo lo que se puede observar es la radiación de la dispersión a un ángulo 2θ. Si expresamos esta intensidad relativa a la intensidad dispersada por un solo electrón en el origen, tenemos la ecuación:

$$I(\mathbf{S}) = \mathbf{F}(\mathbf{S}) \cdot \mathbf{F}^*(\mathbf{S}) = |\mathbf{F}|^2$$

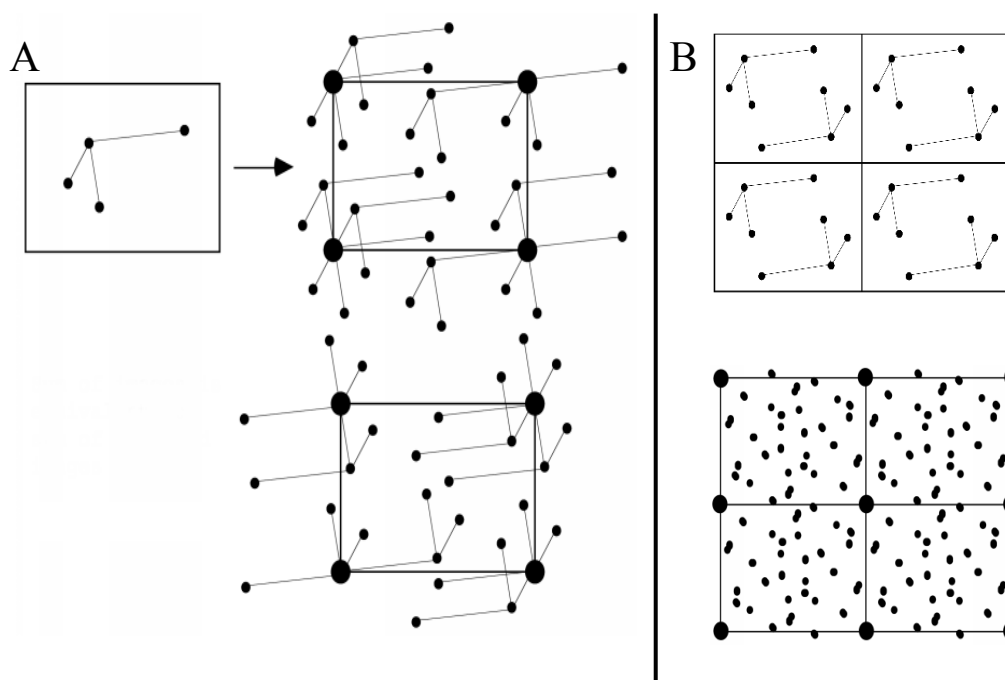
La intensidad es algo real y cuantificable, es directamente proporcional al cuadrado de la amplitud del factor de estructura. De este modo, **|F|** se puede medir experimentalmente, sin embargo el factor correspondiente a la fase ( $\exp i\alpha$ ) no se puede obtener de los datos de difracción. Esto es lo que se llama el problema de la fase. Para solucionar este problema se han desarrollado diferentes métodos indirectos estos son: reemplazamiento isomorfo múltiple o simple (MIR y SIR), reemplazo molecular, difracción anómala a diferente longitud de onda (MAD), y los métodos directos. En el presente trabajo usamos dos métodos, el reemplazamiento isomorfo y el reemplazo molecular. En ambos casos, las intensidades medidas de la difracción del cristal se usan para evaluar la función de Patterson, la cual es una herramienta matemática fundamental en el proceso de la resolución de la estructura.

#### **A1.5.- La función de Patterson**

La función de Patterson, **P(u)** o **P(uvw)** es un sumatorio de Fourier usando como coeficientes las intensidades medidas. De forma que solo se usan datos experimentales y no se necesita conocer la fase para calcularla. Se puede expresar como:

$$P(uvw) = 1/V \sum_{hkl} |\mathbf{F}(hkl)|^2 \cos(2\pi\rho(hu + kv + lw))$$

donde  $u$ ,  $v$  y  $w$  son coordenadas cartesianas normales que describen el espacio de Patterson.



**Fig. A1.3. La función de Patterson.** La función de Patterson proporciona un mapa de vectores interatómicos. Si existe un pico de densidad electrónica para un átomo 1 en una posición  $x_1$  y otro para un átomo 2 en posición  $x_2$ , el mapa de Patterson tendrá picos en las posiciones dadas por  $x_2 - x_1$  y  $x_1 - x_2$ . La intensidad de un pico en el mapa de Patterson es proporcional al producto de las intensidades de los dos picos en el mapa de densidad electrónica. **A**, representación de Patterson correspondiente a una celda unidad con una molécula, se puede pensar en un mapa de Patterson como la suma de imágenes de la molécula, con cada uno de sus átomos desplazados al origen. Como a cada vector le corresponde otro en la dirección contraria, el mismo mapa es también una suma de las imágenes invertidas de la molécula. **B**, mapa de Patterson correspondiente a un cristal con dos moléculas por unidad asimétrica.

Patterson demostró que el sumatorio de Fourier de las intensidades es equivalente al conjunto de todos vectores interatómicos en la molécula, cada uno relacionado con el producto de los números atómicos del respectivo par de átomos. En términos matemáticos se expresa como:

$$P(uvw) = V \int \rho(x,y,z) \rho(x+u, y+v, z+w)$$

El concepto se entiende mejor con un diagrama como el de la (Fig.A1.3.).

El mapa de Patterson resultante es siempre centrosimétrico y posee una simetría mayor que la correspondiente al cristal de proteína. La intensidad de un pico de Patterson es proporcional al producto del número de electrones (número atómico) de los átomos en cada uno de los extremos de los vectores interatómicos. El número de picos de un mapa de Patterson, excluyendo los picos del origen, es igual a  $N(N-1)$ , donde  $N$  es el número de átomos de la molécula.

Resumiendo, la función de Patterson tiene la ventaja de que puede ser calculada directamente de los datos medidos, y aunque no proporciona la estructura de la molécula (como hace la densidad electrónica), genera una función relacionada con la estructura de una forma muy precisa y conocida.

#### **A1.6.- El método de reemplazamiento isomorfo (MIR, SIR)**

El método de reemplazamiento isomorfo múltiple (MIR) consiste en introducir cambios en el cristal que perturben los factores de estructura, con la esperanza de que estas alteraciones nos permitan deducir los posibles valores de fase. Para ello hay que medir datos de diferentes cristales de proteína, un conjunto de datos corresponde a un cristal de proteína nativa, y los otros son derivados de átomos pesados de los cristales nativos. Los cristales se “derivatizan” introduciendo uno o varios átomos de metal de alto peso molecular, como por ejemplo, Hg, Au, Pt, Br, Pb, W, U, o bien alguna de las denominadas tierras raras. Los datos de cada uno de los derivados junto con los del nativo se usan para calcular mapas de diferencia de Patterson, de los cuales se puede obtener la posición de los átomos pesados, para que esto funcione el isomorfismo o similitud entre los cristales es fundamental. El conocimiento de la posición de los átomos pesados junto con las amplitudes medidas del cristal nativo y los derivados, permiten el cálculo de un conjunto de fases iniciales aproximadas. De este modo se puede obtener un mapa de densidad electrónica que sirva para determinar la estructura de la proteína.

##### **A1.6.1.- Preparación de cristales derivados**

La preparación de cristales derivados es un procedimiento sumamente empírico, ya que no todos los cristales toleran los átomos pesados de la misma manera. La técnica

habitual consiste en transferir el cristal de su solución original de *sitting* o *hanging drop* a una solución estabilizada que contenga la concentración apropiada del metal pesado elegido, normalmente esta solución ha de contener los mismos componentes que la solución de cristalización para evitar que los cristales se disuelvan. Hay que controlar fundamentalmente dos variables, la concentración de metal pesado y el tiempo de exposición al mismo. Normalmente se usa un exceso de metal pesado, con la intención de desplazar la reacción hacia el producto (derivativo). Pero hay que tener cuidado, ya que una concentración demasiado alta puede dañar el cristal y disminuir el isomorfismo. Normalmente hay que hacer varios ensayos de concentración y tiempo para dar con la condición ideal, en la cual el cristal no se rompa y además difracte. Además si la presencia de metal pesado es muy alta, éste puede unirse a numerosos sitios, lo que hace más dificultosa la interpretación de la función de Patterson. El tiempo de incubación puede variar desde unos minutos, a horas o incluso días. El tiempo de reacción entre la proteína y el metal hasta alcanzar el equilibrio depende de numerosos factores. Primero, de la difusión del reactivo por los poros del cristal, lo cual depende de la relación entre el tamaño de los poros y los reactivos. Segundo, puede que la proteína necesite un ligero cambio conformacional para poder acomodar el átomo de metal pesado. Finalmente del tiempo que tarde la reacción química en producirse.

El sitio de unión de los átomos pesados a las proteínas depende de la naturaleza del propio átomo pesado. Así, el Hg se une a grupos SH<sup>-</sup> libres, y también a las cadenas laterales de histidinas si el pH del medio no es demasiado bajo. El Pt tiene preferencia por los átomos de azufre de las cadenas laterales de las metioninas, pero también se ha observado en las proximidades de las cadenas laterales de las argininas. El Pb se une preferentemente a las cadenas laterales de residuos glutámico de la superficie de la molécula proteica. En proteínas que contengan Ca<sup>2+</sup> o Mg<sup>2+</sup>, dichos iones pueden ser sustituidos por iones de similar radio atómico pero con un número mayor de electrones.

Hay muchos factores que afectan a la unión de los átomos pesados. Los más importantes son: el pH, la constante dieléctrica, la fuerza iónica y la presencia de agentes quelantes en la solución.

Hay casos en los cuales no es posible utilizar estos métodos, porque la proteína no tiene sitios de unión a los átomos pesados. Cuando pasa esto se puede recurrir a otros métodos alternativos. Uno de estos métodos consiste en modificar químicamente la proteína nativa mediante la unión covalente con un agente que sea capaz de unirse a un átomo pesado o bien que contenga dicho átomo pesado. Otra posibilidad es incorporar



selenometionina en la proteína expresándola en un modelo adecuado que no incorpore metionina.

#### **A1.6.2- Calidad de los datos de rayos-X colectados**

No siempre los cristales obtenidos son apropiados para resolver la estructura tridimensional de una proteína, a veces, aunque a simple vista los cristales parezcan buenos no difractan lo suficiente o bien difractan pero tienen algún problema en su composición que hace que las estadísticas obtenidas tras procesar los cristales no sean buenas.

Los principales criterios para evaluar la calidad de los datos cristalográficos son:

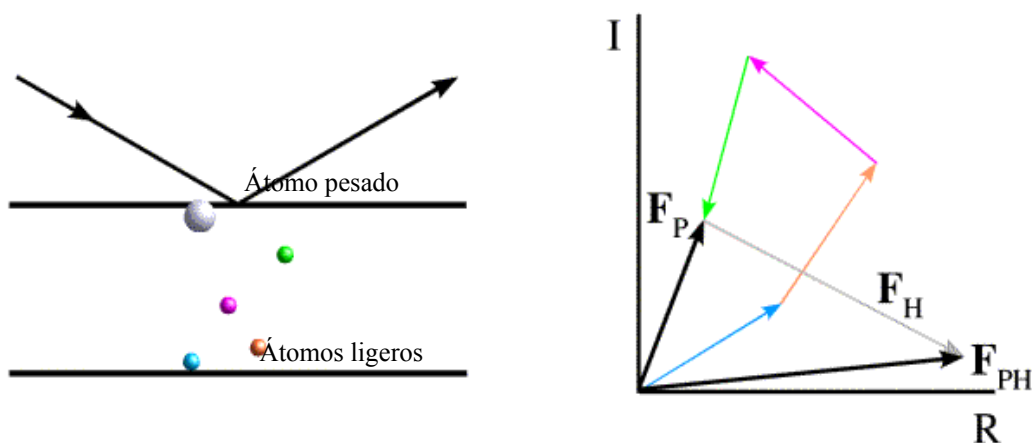
- El número de reflexiones únicas y la fracción de todas las posibles reflexiones teóricas en el conjunto de datos (completitud).
- El número de mediciones (redundancia).
- La consistencia interna de los datos (R-merge), es la diferencia entre los datos experimentales y los calculados.
- El número o fracción de medidas rechazadas de las estadísticas.
- El valor de la intensidad sobre sigma, total y a una resolución de terminada.
- La resolución y la completitud para el último valor de resolución medido (el más alto).

#### **A1.6.3.- Efecto de los átomos pesados en el patrón de difracción**

El método del reemplazamiento isomorfo se basa en la función de Patterson calculada con la diferencia de las amplitudes de los factores de estructura de las proteínas nativa y derivada.

La contribución de los átomos más ligeros tiende a cancelarse entre si, porque difractan con ángulos de fase diferentes (FigA1.4.). Sin embargo, todos los electrones de un átomo pesado difractan en fase unos con otros (FigA1.4.)

Debido a este efecto, los diferentes átomos contribuyen a la intensidad dispersada de forma proporcional al cuadrado del número de electrones que contienen. Por ejemplo,



**FigA1.4. Reemplazamiento isomorfo.** La contribución de los átomos ligeros a la intensidad total tiende a neutralizarse, ya que difractan con ángulos de fase diferente. Sin embargo, los electrones que forman un átomo pesado difractan en fase unos con otros.

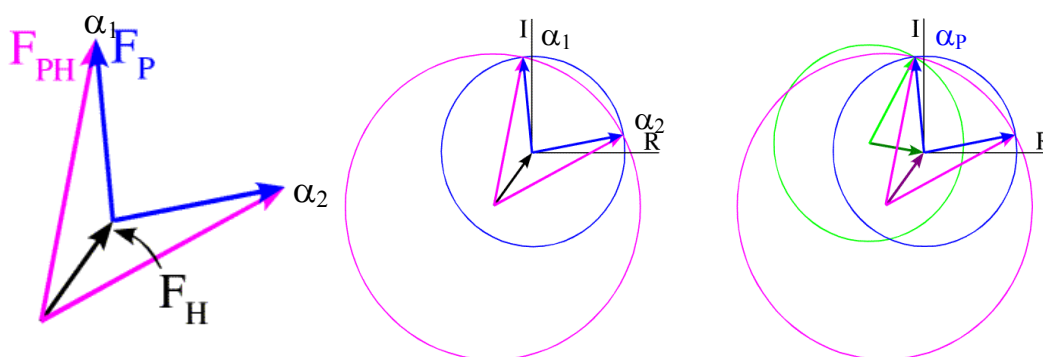
un átomo de uranio, contiene 15 veces más electrones que uno de carbono, por lo tanto su contribución a la intensidad total será equivalente a 225 átomos de carbono. El cambio de intensidad esperada por la introducción de un átomo pesado fue descrito por Crick y Magdoff en 1956. Calcularon que dicho cambio de intensidad es del orden de 0,25 – 0,50 %, para una proteína de tamaño medio, lo que supone un cambio lo suficientemente significativo como para ser detectado.

A la hora de localizar la posición de los átomos pesados unidos al cristal para obtener un conjunto de fases que permitan resolver la estructura de la proteína se recurre nuevamente a métodos de Patterson. Para una proteína normal, el número esperado de picos en un mapa de Patterson es del orden de millones, la gran cantidad de picos hace que el mapa sea ininteligible. Para determinar la posición de un átomo pesado unido a una proteína en un cristal se calcula un mapa de Patterson con unos coeficientes iguales al cuadrado de la diferencia isomorfa. La diferencia isomorfa se define como la diferencia entre las amplitudes de los factores de estructura de las proteínas nativa y derivativa, (las cuales se han medido experimentalmente). Lo cual implica, que si los cristales son lo suficientemente isomorfos, el mapa de Patterson resultante solo mostrará los picos generados por el átomo pesado, mas algo de ruido. Como el número de

átomos pesados unidos al cristal es pequeño, el mapa de Patterson tendrá un número de picos manejable, que se podrá interpretar.

La localización de los átomos pesados a partir de un mapa de Patterson se basa en las relaciones de simetría del grupo espacial del cristal. Se toman las coordenadas  $uvw$  de cada pico del mapa de Patterson y mediante las relaciones de simetría se obtienen las coordenadas  $xyz$  del átomo pesado en el cristal.

El procedimiento geométrico para calcular la fase de la proteína a partir de los factores de estructura de un átomo pesado se llama construcción de Harker (Fig.A1.5.). Se basa en el hecho de que para ambos vectores  $F_P$  y  $F_{PH}$  solo conocemos las amplitudes (la



**Fig.A1.5. La construcción de Harker**, es un truco geométrico para evaluar el ángulo de la fase correspondiente a una proteína  $\alpha_p$ . Para construirla se dibuja un círculo cuyo radio es igual a la amplitud del factor de estructura de la proteína ( $|F_P|$ ), desde el centro de este círculo se dibuja un vector equivalente al negativo del factor de estructura del átomo pesado ( $|F_H|$ ). A continuación se dibuja un círculo de radio igual a la amplitud del factor de estructura de la proteína derivativa ( $|F_{PH}|$ ) colocado al final del vector  $|F_H|$ . Los dos círculos se cortan en dos puntos que corresponden a los dos posibles valores para los ángulos de la fase de la proteína. Esta construcción cuenta con el problema de la ambigüedad de fase, la cual se puede solucionar añadiendo un segundo derivativo.

longitud de los vectores) pero no las fases (la orientación de dichos vectores). Esto quiere decir que podrían estar orientados hacia cualquier dirección dentro de los círculos. La posición relativa de un círculo con respecto del otro viene fijada por el vector  $F_H$ , el cual es conocido. De este modo, solo los puntos en que ambos círculos se interceptan proporcionarían orientaciones para  $F_P$  y  $F_{PH}$  que satisfagan la diferencia vectorial:  $F_P = F_{PH} - F_H$ .

Hay que destacar que la construcción de Harker da dos posibles valores para la fase de la proteína. Esto se puede solucionar trabajando con un segundo derivativo, de este modo solo uno de los valores de la fase satisfará la construcción geométrica para los dos derivativos al mismo tiempo, este método se conoce como MIR (*múltiple isomorphous replacement*) en contraste al método que sólo utiliza un derivativo (*single isomorphous replacement*).

### A1.7.- El método MAD

Los cambios que se provocan en la intensidad de la difracción al introducir átomos pesados en los cristales de proteína se pueden considerar como modificaciones químicas de la difracción. De modo análogo, se pueden provocar cambios en las intensidades de la difracción modificando las propiedades físicas de los átomos. Así, si la radiación X incidente tiene una frecuencia próxima a la frecuencia natural de oscilación de los electrones de un determinado átomo, se produce la denominada *dispersión anómala*, que modifica el factor de dispersión atómico,  $f_j$ , de tal modo que su expresión se ve modificada con dos términos  $f'$  y  $f''$ , que son las componentes real e imaginaria, respectivamente, de la fracción anómala del factor de dispersión. Así, el factor de dispersión para los átomos que se comportan anómalamente viene dado por la expresión:

$$f = f^0 + f' + i f''$$

Las variaciones de  $f'$  y  $f''$  frente a la energía de los rayos X se pueden calcular usando consideraciones teóricas. Para los valores de energía de los rayos X en los que existe resonancia, el valor de  $f''$  aumenta mucho, al mismo tiempo que el valor de  $f'$  decrece. Esto es muy importante en cristalografía, ya que muchos de los átomos pesados utilizados tienen picos de absorción a energías (longitudes de onda) que se pueden obtener de manera fácil utilizando radiación sincrotrón. De este modo, los datos de difracción colectados en estas condiciones tendrán una componente normal, debida a los átomos ligeros (nitrógeno, carbono e hidrógeno), y una componente anómala, debida a los átomos pesados, que producirá una alteración en la fase global de cada reflexión. Esto se traducirá en un cambio de intensidad entre las llamadas reflexiones de Friedel (que son parejas de reflexiones que en condiciones normales deberían tener la misma amplitud e idénticas fases, pero signos opuestos). El cambio de intensidad detectable entre la parejas de Friedel recibe el nombre de *difracción anómala*.

El método MAD implica la medida de los datos de difracción de un cristal de la proteína (que contenga un dispersor anómalo fuerte) usando radiaciones de distintas energías: la que maximiza  $f''$ , la que minimiza  $f'$  y una energía lejana de estas dos. Combinando estos conjuntos de datos de difracción, y analizando las diferencias entre ellos, es posible calcular la distribución de amplitudes y fases que generan los

dispersores anómalos. El uso de las fases generadas por estos dispersores anómalos, permite calcular la densidad electrónica para toda la proteína.

La principal ventaja del método MAD con respecto del MIR es que como los datos se recogen de un único cristal, desaparecen los problemas de falta de isomorfismo que pueden darse de forma habitual en los experimentos de MIR.

La desventaja es que para llevar a cabo experimentos de MAD es imprescindible acudir a los centros de radiación sincrotrón, en donde es posible sintonizar diferentes longitudes de onda en función del experimento que se quiera llevar a cabo.

### **A1.8.- Método de reemplazo molecular**

Los miembros de una familia de proteínas tienen patrones de plegamiento similares y a menudo comparten una alta homología en la secuencia. Si se conoce la estructura de uno de los miembros de la familia, la estructura de los demás se puede obtener por reemplazo molecular. La similitud estructural se puede basar en una alta similitud en la secuencia, o bien en que las proteínas tengan idéntica función, o que sean mutantes de la misma proteína.

El punto de partida es un conjunto de datos nativos y un modelo tridimensional de la proteína relacionada. La posición y la orientación de la molécula de proteína en la celda unidad no se conoce *a priori*, y ha de ser determinada. Para llevar a cabo este proceso hay que evaluar dos funciones de Patterson. Primero se obtiene un mapa de Patterson experimental a partir de los datos cristalográficos de la molécula desconocida. El mapa de Patterson calculado se obtiene a partir del modelo tridimensional de la molécula conocida, cuya orientación y posición en la celda unidad se desconoce. El mapa de Patterson calculado es rotado y trasladado varias veces. A cada paso, los mapas calculado y experimental se comparan y se evalúa el grado de coincidencia entre ellos. La solución se encuentra cuando la comparación muestra un alto grado de correlación entre ambos mapas.

La resolución de una estructura por reemplazo molecular es un problema con seis variables, tres ángulos para definir la rotación y tres coordenadas para definir la traslación. La rotación se lleva a cabo primero porque la búsqueda de la traslación depende mucho de haber hallado la correcta orientación del modelo.

#### **A1.8.1.- La función de rotación**

Para explicar la función de rotación de una manera simple, consideremos una molécula con sólo tres átomos. Esta molécula, A, tiene una estructura conocida,  $\rho_A(\mathbf{r})$ , y una orientación inicial arbitraria conocida. La función de Patterson  $P_A(\mathbf{u})$ , para esta molécula se puede calcular. Supongamos que tenemos datos de difracción de rayos X medidos para una molécula desconocida B, la cual es muy parecida a la molécula A. Lo primero que hay que hacer es conocer la orientación de la molécula B en la celda unidad.

La función de Patterson calculada para la molécula conocida A, en su orientación inicial, dará lugar a un conjunto de vectores que conectan cada par de átomos de la molécula. A partir de las intensidades medidas de la molécula B, se obtiene una función de Patterson experimental,  $P_B(\mathbf{u})$ , con sus correspondientes vectores. Estos vectores se denominan *auto vectores* y el mapa obtenido *auto Patterson*.

El proceso consiste en rotar la molécula conocida A en pequeños pasos angulares, a cada paso la densidad electrónica viene dada por  $\rho_A(\mathbf{C}\mathbf{r})$ , donde  $\mathbf{C}$  es la matriz de rotación. La función de Patterson de la molécula rotada,  $P_A(\mathbf{C}\mathbf{u})$ , se puede entonces calcular y comparar con la función de Patterson experimental de la molécula B,  $P_B(\mathbf{u})$ . El grado de coincidencia entre  $P_B(\mathbf{u})$  y  $P_A(\mathbf{C}\mathbf{u})$ , se mide mediante la función:

$$R(\Omega) = \int P_B(\mathbf{u})P_A(\mathbf{C} \bullet \mathbf{u}) d\mathbf{u}$$

Donde  $\Omega$  representa los tres valores angulares que definen la orientación de la molécula. Cuando la función de correlación  $R(\Omega)$  tiene un valor alto, quiere decir que los mapas de Patterson coinciden y por tanto que la orientación correspondiente al modelo proporciona la orientación de la molécula en la celda.

### **A1.8.2.- La función de traslación**

La función de rotación se basa en la rotación de una función de Patterson alrededor de un eje, en este punto no se usa traslación. Sin embargo, para la resolución final de la estructura por el método de reemplazo molecular, se debe determinar la traslación de la molécula ( o subunidad ), por superimposición de una molécula sobre otra, de orientación conocida, en el espacio real. Una vez se haya determinado correctamente la orientación por medio de la función de rotación.

La búsqueda de la traslación consiste en encontrar la posición de la molécula desconocida, relativa a los elementos de simetría cristalográficos. El proceso comienza calculando un mapa de Patterson experimental a partir de los datos de difracción de rayos X obtenidos de la molécula B. Es lo que se llama un *Patterson cruzado* porque contiene vectores *cruzados* (*cross vectors*), son vectores que conectan átomos entre dos moléculas diferentes. Entonces, se calcula el Patterson para el modelo correctamente orientado en una posición inicial arbitraria, (normalmente con el centro de masas de la molécula). El problema reside en que no se sabe cuánto hay que trasladar la molécula del modelo hasta alcanzar la posición correcta en el cristal. El modelo se va moviendo paso a paso y en cada posición se calcula la función de Patterson y se compara con el Patterson experimental. La correlación entre las funciones de Patterson calculada y experimental ( lo que indica el grado de coincidencia ) se expresa así:

$$T(\mathbf{t}) = \int P_B(\mathbf{u}) P_{AA'}(\mathbf{C} \bullet \mathbf{u}, \mathbf{t}) d\mathbf{u}$$

donde el vector  $\mathbf{t}$  representa al vector de traslación.

#### **A1.9.- Interpretación de los mapas y construcción del modelo**

Una vez que se ha obtenido un conjunto de fases y se ha calculado un primer mapa de densidad electrónica, se puede comenzar a construir un modelo de la proteína consistente con dicha densidad electrónica.

Este procedimiento se lleva a cabo usando ordenadores con alta capacidad gráfica. Hay numerosos programas para interpretar estos resultados, uno de los más utilizados es O, desarrollado por T.A. Jones (Jones TA 1991). La construcción del modelo dentro de la densidad electrónica es un proceso difícil, las primeras fases obtenidas son aproximadas y el primer mapa de densidad electrónica calculado con dichas fases suele tener mucho ruido. En este punto, el investigador ha de hacer su propia interpretación de los resultados y ha de tener mucho cuidado para evitar introducir errores. El proceso se lleva a cabo siguiendo los siguientes pasos:

- Determinar los límites de la molécula. Esto es, identificar la porción de la densidad electrónica que pertenece a una molécula. Esto, a veces, resulta sencillo, pero en algunas ocasiones puede ser complicado. Un posible error

puede ser incluir en el modelo un segmento de una molécula cercana relacionada por la simetría del grupo espacial o bien por simetría no cristalográfica (NCS). Esta simetría no cristalográfica es la que posee la proteína por formar un dímero, un trímero u otra forma de oligomerización, esta simetría es local y no se extiende a toda la red cristalina.

- Determinar el plegamiento de la proteína. Esto significa: identificar los elementos de la estructura secundaria de la proteína ( $\alpha$ -hélices y láminas- $\beta$ ), de la estructura terciaria y los diferentes dominios estructurales. Encontrar como se conectan los elementos de estructura secundaria observados y los dominios. Identificar dónde se localizan las cadenas laterales o hacia donde apuntan. Y encontrar la dirección de la cadena polipeptídica.
- Asignar las posiciones de los carbonos  $\alpha$  en el modelo.
- Construir la cadena principal (N, C, O,  $C\alpha$ ).
- Generar las cadenas laterales siguiendo el mapa de densidad electrónica.
- Finalmente hay que comprobar y corregir los posibles errores introducidos.

#### A1.10.- Refinamiento cristalográfico

El primer modelo obtenido es solo aproximado. Por lo tanto, el conjunto de factores de estructura calculados,  $\{F_{hkl}\}_{calc}$ , tendrá poca coincidencia con el conjunto de factores de estructura medidos,  $\{F_{hkl}\}_{obs}$ . El grado de coincidencia entre los factores de estructura observados y calculados se representa por el factor-R, y se representa así:

$$R (\%) = \frac{\sum_{hkl} ||F_{obs}| - k|F_{calc}||}{\sum_{hkl} |F_{obs}|} \times 100$$

así, para un modelo inicial el valor de R suele ser alrededor del 50%.

Se conoce como refinamiento al proceso de ajustar el modelo hasta encontrar una alta coincidencia entre los factores de estructura calculados y observados. Se han desarrollado numerosos métodos para llevarlo a cabo, si se aplican bien, pueden bajar el factor-R de un modo sustancial, consiguiendo valores en el rango de 10 a 20% o incluso menores. El ajuste del modelo consiste en cambiar los parámetros posicionales y los factores de temperatura para todos los átomos de la estructura, excepto los hidrógenos.



Como los átomos de hidrógeno contienen un único electrón, su influencia en la difracción de los rayos X es muy baja, y por eso no se suelen tener en cuenta en el proceso de determinación de la estructura.

Para evaluar la calidad del proceso de refinamiento cristalográfico se calcula un parámetro llamado factor-R libre,  $R_{\text{libre}}$ . Es un parámetro más fiable para comprobar el refinamiento. Para calcularlo, los datos medidos se dividen en dos grupos, el grupo de trabajo, que contiene la mayoría de los datos, y el grupo de evaluación, que contiene alrededor del 10% de las intensidades medidas. El refinamiento se lleva a cabo con los datos del grupo de trabajo y el factor-R se calcula con el grupo de evaluación ( $R_{\text{libre}}$ ). El  $R_{\text{libre}}$ , al contrario del factor-R cristalográfico, tiene la ventaja de estar libre del ruido que introduce el modelo, por lo tanto es un indicador mejor de la calidad del mismo. Si el refinamiento funciona tanto el  $R_{\text{libre}}$  como el factor-R han de disminuir a la vez que el modelo mejora.

#### **A1.10.1.- Método de los mínimos cuadrados**

La mayoría de las técnicas de refinamiento en cristalografía de proteínas están basadas en el principio de los mínimos cuadrados. El refinamiento por mínimos cuadrados es un proceso repetitivo. Los parámetros a refinar van cambiando en cada ciclo hacia su valor final. Normalmente son necesarios un buen número de ciclos hasta que los cambios en los parámetros son lo suficientemente pequeños, entonces se considera que hay convergencia hasta un conjunto final de parámetros. El rango de convergencia es la distancia máxima de los átomos para moverse a su posición final. Si éstos están muy lejos es probable que la función a minimizar se encuentre atrapada en un mínimo local que no corresponde con el correcto. El progreso del refinamiento se puede monitorizar calculando el factor-R después de cada ciclo.

$$Q = \sum_{hkl} w(hkl) (|F_{\text{obs}}(hkl)| - |F_{\text{calc}}(hkl)|)^2$$

$w$  es el valor dado a una observación. El mínimo de  $Q$  se encuentra cuando el primer derivativo con respecto a cada una de las variables es igual a cero:

$$\partial Q / \partial u_j = 0$$

Esto produce un sistema de ecuaciones que hay que resolver para encontrar los cambios en los parámetros, que se aplicaran a los parámetros  $u_j$ . El proceso es iterativo, hay que calcular nuevos valores de  $|F_{\text{calc}}(hkl, u_j)|$  y sus derivativos con respecto a los  $u_j$ 's para cada ciclo.

#### **A1.10.2.- El factor de temperatura**

Los átomos de un cristal no son estáticos, en realidad están vibrando alrededor de su posición de equilibrio, es lo que se llama vibración térmica. Así, el haz de rayos X a través de un cristal no encuentra a todos los átomos exactamente en la misma posición, en las sucesivas celdas unidad, sino ligeramente desplazados unos con respecto a otros. El efecto es como si el haz de rayos X encontrara un átomo desplazándose en una posición fija. Cuanto más fuerte es la vibración térmica, más grande es el desplazamiento de la densidad electrónica. Este efecto disminuye la intensidad de los rayos X dispersados, especialmente los que tienen mayores ángulos de dispersión,  $\theta$ , esto es los de mayor resolución. Este efecto se regula multiplicando los factores de dispersión por un factor dependiente de la temperatura, llamado factor de temperatura, y se simboliza como B.

#### **A1.10.3.- Dinámica molecular**

La mayoría de los métodos de refinamiento se basan en el principio de los mínimos cuadrados. En estos métodos, la función a minimizar sigue una línea descendente hacia su valor mínimo, y si el modelo de partida no es muy diferente de la estructura real de la proteína, el refinamiento converge fácilmente hacia la solución correcta. Sin embargo, si la distancia entre los átomos del modelo y de la estructura real es bastante grande, el refinamiento se puede quedar bloqueado en un mínimo local en lugar de alcanzar el mínimo global o verdadero. Para evitar esta situación se requiere una técnica capaz de ir en ambas direcciones, subiendo y bajando, para sobrepasar las barreras de la función Q. Esta técnica se denomina “dinámica molecular (Brünger AT 1987).

La idea principal de este método se basa en simular el comportamiento dinámico de las moléculas para sobrepasar las barreras de energía potencial y alcanzar el mínimo global. Si la temperatura del sistema es elevada, muchos átomos tendrán valores de energía cinética altos y podrán sobrepasar altas barreras energéticas. Cuando el sistema

es enfriado de forma lenta, se aproxima al valor de energía mínimo. Este método se denomina “anillamiento simulado” (“simulated annealing”). Un ejemplo de proceso con alta barrera energética es la inversión del plano de un péptido. Esto es posible hacerlo con dinámica molecular pero no con otras técnicas de refinamiento. Esta técnica tiene un elevado radio de convergencia.

La función a minimizar se describe como:

$$E_{\text{total}} = E_i + E_x$$

donde  $E_i$  es la energía interna y  $E_x$  es un termino de pseudoenergía que contiene los datos cristalográficos. La energía interna incluye términos relativos a la estereoquímica de la molécula y a los contactos de los átomos en la proteína plegada.

Se han desarrollado varios programas para llevar a cabo esta técnica, el más utilizado es un paquete de programas llamado CNS (Brünger AT 1998). El proceso de refinamiento cristalográfico usando CNS es el siguiente. Los valores de los términos cristalográficos y los de la energía interna de la función son calculados. Se lleva a cabo un proceso de minimización de la energía del modelo. En este punto se usa dinámica molecular calentando y enfriando, y se origina un nuevo modelo. Éste se vuelve a someter a minimización de la energía y se usa para calcular un nuevo mapa de densidad electrónica. En pasos sucesivos se llevará a cabo también refinamiento de los parámetros térmicos (factor B).

El modelo refinado junto con el mapa se evalúan usando ordenadores con programas de gráficos (O). Si es necesario se hacen ajustes manuales, es lo que se denomina reconstrucción del modelo. Para refinar la estructura se llevan a cabo sucesivos pasos de CNS, seguidos de reconstrucción manual.

#### **A1.10.4.- Mapas de densidad electrónica**

A cada ciclo de refinamiento se obtiene un modelo mejorado y se calculan mapas de densidad electrónica. La ecuación de la densidad electrónica es la siguiente:

$$\rho(x, y, z) = 1/V \sum_{hkl} |F_{hkl}| \exp \{-2\pi i (hx + ky + lz) + i\Phi_{hkl}\}$$

Los dos tipos de mapas de densidad electrónica más utilizados son los mapas de diferencia de Fourier  $F_o - F_c$  y  $2F_o - F_c$ .

El mapa  $F_o - F_c$  tiene los coeficientes  $|F_o| - |F_c|$  y las fases calculadas del modelo. Muestra rasgos del mapa de densidad electrónica que no han sido tenidos en cuenta por el modelo. Los átomos que faltan aparecen como densidad electrónica positiva, mientras que los átomos que sobran se ven como densidad negativa. La altura de los picos es la mitad que un mapa normal. Muestra pequeños cambios en la densidad electrónica, como pérdida o ganancia de una molécula de agua.

El mapa  $2F_o - F_c$  se calcula con los coeficientes  $2|F_o| - |F_c| = 2(|F_o| - |F_c| + |F_c|)$  y las fases calculadas del modelo. Muestra la estructura del modelo mas algunos rasgos que no han sido tenidos en cuenta por éste. Hay que tener en cuenta que las fases obtenidas del modelo siempre generan un poco de ruido.

### A1.11. Bibliografia

- Brünger AT, K. J., Karplus M (1987). "Crystallographic R Factor refinement by molecular dynamics." Science **235**: 458-60.
- Brünger AT, P. D. A., G. Marius Clore, Warren L. DeLano, Piet Gros, Ralf W. Grosse-Kunstleve, Jian-Sheng Jiang, John Kuszewski, Michael Nilges, Navraj S. Pannu, Randy J. Read, Luke M. Rice, Thomas Simonson, and Gregory L. Warren. (1998). "Crystallography & NMR System: A New Software Suite for Macromolecular Structure Determination." Acta Cryst. D **54**: 905-921.
- Jones TA, Z. J., Cowan SW, Kjeldgaard. (1991). "Improved methods for building protein models in electron density maps and the location of errors in these models." Acta Cryst. A **47**: 110-119.



## **ABREVIATURAS**





## **ABREVIATURAS**

- **Å:** Ångstrom
- **ADN:** Ácido desoxirribonucleico
- **ADNc:** Ácido desoxirribonucleico codificante
- **Ala:** Alanina
- **APC:** Células presentadoras de antígenos (*Antigen-presenting cells*)
- **Asn:** Asparagina
- **Asp:** Ácido aspártico
- **BCECF:** (2',7'-bis(carboxyethyl)-5(6)-carboxyfluorescein)
- **BSA:** (*Bovine Serum Albumin*)
- **CHO:** (*Chinese Hamster Ovary*)
- **CRD:** dominios de reconocimiento de carbohidratos (*Carbohydrate recognition domain*)
- **Cα:** Carbono alfa
- **D:** Dominio
- **DC-SIGN:** (*dendritic cell-specific ICAM-grabbing nonintegrin*)
- **DMEM:** (*Dulbecco's Modified Eagle's Medium*)
- **EDTA:** ácido etilendiaminotetraacético (*ethylenediamine tetraacetic acid*)
- **ELISA:** (*enzyme-linked immunosorbent assay*)
- **ESRF:** (*European synchrotron radiation facility*)
- **FPLC:** (*Fast Protein Liquid Chromatography*)
- **Fuc:** Fucosa
- **Gal:** Galactosa
- **glc:** Glucosilación
- **GlcNAc:** N-acetilgalactosamina
- **GlcNAc:** N-acetilglucosamina
- **Glu:** Ácido glutámico
- **Gly:** Glicina
- **GPCRs:** (*G protein-coupled receptors*)
- **HEPES:** (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
- **HRP:** (*horseradish peroxidase*)
- **HRV:** Rinovirus humano (*Human rhinovirus*)

- **IC:** (ver ICAM)
- **ICAM:** (*Intercellular Adhesion Molecule*)
- **Ig:** Inmunoglobulina
- **IgSF:** Superfamilia de las Inmunoglobulinas
- **kDa:** kilodalton
- **l:** litro
- **Leu:** Leucina
- **LFA:** (*Lymphocyte Function-Associated Antigen*)
- **M:** Molar
- **MAdCAM-1:** (*mucosal addressin cell adhesion molecule-1*)
- **Man:** Manosa
- **MBP:** Proteína de unión a manosa (*Manose binding protein*)
- **MHC:** Complejo principal de histocompatibilidad (*major histocompatibility complex*)
- **MIDAS:** (*metal ion-dependent adhesion site*)
- **MIR:** Reemplazamiento isomorfo multiple
- **ml:** mililitro
- **mM:** milimolar
- **nm:** Nanómetro
- **nm:** Nanómetro
- **°C:** Grado centígrado
- **OPD:** (*orthophenylene-diamine*)
- **PBS:** (*Phosphate Buffered Saline*)
- **PCR:** Reacción en cadena de la polimerasa (*Polymerase Chain Reaction*)
- **PEG:** Polietilenglicol
- **Phe:** Fenilalanina
- **pI:** Punto isoeléctrico
- **Pro:** Prolina
- **PSGL:** (*P-selectin glycoprotein ligand*)
- **Rpm:** Revoluciones por minuto
- **Ser:** Serina
- **TCR:** Receptor de la célula T (*T-cell receptor*)
- **Thr:** Treonina

- **Tyr:** Tirosina
- **Val:** Valina
- **VCAM:** (*vascular cell adhesion molecule*)
- **VIH:** Virus de inmunodeficiencia humana
- **VLA:** (*very late activation antigen*)
- **µg:** Microgramo
- **µm:** Micrómetro



## **INDICE**

RESUMEN.....	1
1. INTRODUCCIÓN.....	3
1.1. Adhesión celular.....	5
1.1.1. Adhesión dinámica, migración celular.....	5
1.1.2. Adhesión transitoria.....	8
1.1.3. Adhesión estática.....	11
1.2. Moléculas de adhesión celular.....	13
1.2.1. La subfamilia de los ICAMs.....	16
1.2.2. ICAM-1.....	17
1.2.3. ICAM-2 .....	17
1.2.4. ICAM-3.....	19
1.2.5. ICAM-5.....	19
1.3. N-glicosilaciones en los miembros de la subfamilia de los ICAMs.....	21
1.4. Estructura, biosíntesis y función de los carbohidratos.....	22
1.5. Lectinas.....	24
1.5.1. Lectinas de tipo C y sus ligandos.....	25
1.5.2. DC-SIGN.....	25
2. OBJETIVOS.....	29
3. MATERIALES Y MÉTODOS.....	33
3.1. Plásmidos y mutagénesis.....	35
3.2. Preparación de proteínas de fusión con Fc solubles.....	37
3.3. Preparación de DC-SIGN.....	37
3.4. Unión de ICAMs a DC-SIGN.....	38
3.5. Tratamiento de ICAMs-Fc con endoglicosidasa H y análisis de las proteínas.....	38
3.6. Expresión de ICAMs en la superficie celular y ensayos de unión con anticuerpos, LFA-1 y HRV.....	39

3.7. Resolución de la estructura tridimensional de IC5-4D.....	41
3.7.1. Preparación y purificación de IC5-4D.....	41
3.7.2. Concentración y cristalización de IC5-4D.....	41
3.7.3. Colección y procesamiento de datos cristalográficos.....	43
3.7.4. Resolución y refinamiento de la estructura.....	43
4. RESULTADOS.....	47
4.1. Contribución de los glicanos en el plegamiento y función de los ICAMs.....	49
4.1.1. Un motivo proteína-glicano conservado en el primer dominio de los ICAMs.....	50
4.1.2. Contribución de los glicanos al plegamiento y expresión del dominio N-terminal de ICAM-1 aislado.....	53
4.1.3. Importancia de las N-glicosilaciones en la expresión de ICAM-2 y en el reconocimiento de LFA-1.....	55
4.2. Contribución de las N-glicosilaciones al reconocimiento de ICAMs por la lectina DC-SIGN.....	56
4.2.1. Preparación del dominio de lectina de DC-SIGN.....	56
4.2.2. Glicosilación en ICAMs y su contribución al reconocimiento por DC-SIGN.....	57
4.3. Estructura de ICAM-5.....	62
4.3.1. Preparación de fragmentos de ICAM-5 con los dominios extracelulares implicados en adhesión homofílica.....	62
4.3.2. Cristalización de la proteína.....	64
4.3.3. Colección y procesamiento de datos cristalográficos.....	65
4.3.4. Determinación de la estructura de la forma X3.....	66
4.3.5. Determinación de la estructuras de la formas X1 y X2.....	72
4.3.6. Validación del proceso de resolución de las estructuras.....	74
4.3.7. Empaquetamiento de las moléculas dentro de la celda unidad en las tres formas cristalinas resueltas.....	75
4.3.8. Estudio de la movilidad entre los dominios en las tres formas cristalinas.....	77
4.3.9. Análisis comparativo de la estructura de ICAM-5 con otros miembros de la subfamilia de los ICAMs.....	79

5. DISCUSIÓN.....	85
5.1. Contribución de las N-glicosilaciones a las propiedades funcionales de los ICAMs.....	87
5.2. Importancia de las N-glicosilaciones en la unión de ICAMs a la lectina DC-SIGN.....	89
5.3. Estructura de ICAM-5 y modelos de interacción homofílica.....	91
5.3.1. Preparación de la proteína.....	91
5.3.2. Variabilidad en la cristalización de IC5-4D y resolución de la estructura.....	92
5.3.3. Estructura de los cuatro dominios N-terminales de ICAM-5.....	93
5.3.4. Modelo para la interacción homofílica en ICAM-5.....	94
6. CONCLUSIONES.....	99
7. BIBLIOGRAFÍA.....	103
ANEXO 1.....	111
ABREVIATURAS.....	135

